

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Next Generation Lentiviral Vectors for the Gene Therapy of Hemoglobinopathies

**Permalink**

<https://escholarship.org/uc/item/5vg9970x>

**Author**

Morgan, Richard

**Publication Date**

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

**Next Generation Lentiviral Vectors for the Gene Therapy of Hemoglobinopathies**

A dissertation submitted in partial satisfaction of the requirements for the  
degree Doctor of Philosophy in Molecular & Medical Pharmacology

by

Richard Morgan

2019

© Copyright by  
Richard Morgan  
2019

# **ABSTRACT OF THE DISSERTATION**

## **Next Generation Lentiviral Vectors for the Gene Therapy of Hemoglobinopathies**

by

Richard Morgan

Doctor of Philosophy in Molecular & Medical Pharmacology

University of California, Los Angeles, 2019

Professor Donald B. Kohn, Chair

### **Abstract**

We set out to improve upon current  $\beta$ -globin expressing lentiviral vectors (LV) designs by rationally reengineering human genomic sequences through the removal and/or addition of known elements. Although this approach generated improved LV designs, it is a low-throughput model for LV development, where large sections of DNA containing roughly-defined elements are simply inserted and/or removed from LVs and tested against a litany of criteria until a combination of well-performing elements are found (typically occurring over a span of years). To speed up LV development, we concurrently developed an advanced LV-engineering technology that can be broadly applied to the design of any gene therapy LV that requires high-level and lineage specific expression. Our Lentiviral Vector based, Massively Parallel Reporter Assay (LV-MPRA) allowed us to map the precise boundaries of lineage specific enhancers within large genomic regions (>16kb). The enhancer maps generated guided the assembly of novel enhancer combinations that when placed in a globin expression LV

conferred robust transgene expression within a minimal DNA footprint. While our original “intelligent” design approach generated LV designs with improved performance (when compared to antecedent clinical LVs), vectors designed using the LV-MPRA based enhancer mapping approach possessed similar improvements in performance and were generated in only a fraction of the time (~3 months vs ~3 years). Moreover, LVs generated by either approach corrected the “Townes” mouse model of Sickle Cell Disease. The yields of this research offer improved “next generation” LVs for the gene therapy of hemoglobinopathies.

The dissertation of Richard A. Morgan is approved.

Samson A Chow

Dinesh Subba Rao

Stephen Smale

Donald Barry Kohn, Committee Chair

University of California, Los Angeles

2019

This thesis is dedicated to my mom, Belinda L. Gilchrist, for always emphasizing education above all else, for providing a role-model of what a life-long learner is, and for igniting a desire to realize great dreams.

## Table of Contents

<b>Abstract of Dissertation</b>	<b>iii</b>
<b>Committee Page</b>	<b>iv</b>
<b>Dedication Page</b>	<b>v</b>
<b>Table of Contents</b>	<b>vi</b>
<b>List of Tables</b>	<b>ix</b>
<b>List of Figures</b>	<b>x</b>
<b>Acknowledgments</b>	<b>xiii</b>
<b>Vita</b>	<b>xvi</b>
<b>Chapter 1 Hematopoietic Stem Cell Gene Therapy –Progress and Lessons Learned</b>	<b>1</b>
Abstract	<b>2</b>
Introduction	<b>2</b>
HSCs for Gene Therapy	<b>4</b>
Vector Choice and Design	<b>6</b>
Safety Issues	<b>7</b>
Methods to optimize Vector Delivery	<b>10</b>
Advances in Vector Design	<b>11</b>
Gene Expression from Lentiviral Vectors	<b>13</b>
Lentiviral Vectors in Clinical Trials	<b>15</b>
X-linked Adrenoleukodystrophy (X-ALD) HSC Gene Therapy	<b>16</b>
Gene Editing	<b>18</b>
Gene Editing Strategies	<b>20</b>
Quiescence	<b>23</b>
Methods of Delivery	<b>25</b>
Gene Editing of HSCs for Clinical Applications	<b>26</b>
Challenges to Clinical Application of HSCs Gene Therapy	<b>26</b>

Lessons Learned	31
Conclusion	37
Tables, Figures and Legends	38
References	48
<b>Chapter 2</b> $\beta$ -Globin Locus Control Region Core Sequences Driving Expression of Anti-Sickling Globin Ameliorates Disease Phenotype in a Mouse Model of Sickle Cell Disease	<b>83</b>
Abstract	<b>84</b>
Introduction	<b>84</b>
Results	<b>87</b>
Proviral Length Influences Packaging and Transduction Efficiency	<b>87</b>
Redefining the Putative Boundaries of HS Elements	<b>90</b>
<i>In vivo</i> Analysis of Peripheral Blood from SCD Mouse Model	<b>93</b>
<i>In vivo</i> Analysis of Bone Marrow $\beta$ AS3-Globin Expression from SCD Mouse Model	<b>96</b>
Incorporating ENCODE Core Sequences into a Shorter $\beta$ -LV Expression Cassette	<b>98</b>
Discussion	<b>98</b>
Materials and Methods:	<b>102</b>
Tables, Figures and Legends	<b>109</b>
Supplemental Results and Figures	<b>126</b>
References	<b>141</b>
<b>Chapter 3</b> Next Generation $\beta$ -globin-Expressing Lentiviral Vectors Created Using a Massively Parallel Reporter Assay – Applying Rapid Dissection of a Locus Control Region to Accelerate Gene Therapy Vector Development	<b>148</b>
Abstract	<b>149</b>
Introduction	<b>149</b>
Results	<b>152</b>
LV-MPRA Identifies Sequence Intrinsic Enhancers of B-Globin Locus Control Region	<b>152</b>

LV-MPRA Guided Therapeutic Vector Design and Characterization	<b>155</b>
In Vivo Characterization of LV-MPRA Based Therapeutic Vectors in SCD Mouse Model	<b>159</b>
Discussion	<b>161</b>
Materials and Methods:	<b>164</b>
Tables, Figures and Legends	<b>175</b>
References	<b>199</b>

## List of Tables

### Chapter 1

Table 1	Genetic diseases of blood cells and the transplantation modalities that have been applied clinically as therapies or are in pre-clinical development	38
---------	--	----

### Chapter 2

Table 1	<i>In vitro</i> VCN of gene modified $\beta$ S/ $\beta$ S Lin(-) BM Cells before transplant	109
Table 2	VCNs of bulk bone marrow cells used to seed methylcellulose based CFU assay	110

### Chapter 3

Tables 1	Comparison of normalized expression between culture conditions.	175
Tables 2	Sequences of primers used in this study.	176

## List of Figures

### Chapter 1

Figure 1	Overview of targets for gene therapy	39
Figure 2	Autologous hematopoietic stem cell transplantation combined with gene addition or editing	42
Figure 3	Summary of Gene Editing Pathways	45

### Chapter 2

Figure 1	Reduction of proviral length increases titer and infectivity	111
Figure 2	Approach for length reduction and <i>In vitro</i> characterization of reduced length vectors	114
Figure 3	Outline of in vivo experiments in a sickle cell disease (SCD) mouse model	117
Figure 4	In vivo analysis of hematologic correction of SCD mouse model	119
Figure 5	In vivo analysis of bone marrow $\beta$ AS3-globin expression from SCD mouse model	121
Figure 6	Further length reduction offers improved vector performance	124
Supplemental Figure 1	Assessment of requirement for WPRE	131
Supplemental Figure 2	Evaluation of larger IVS2 sequence deletion	133
Supplemental Figure 3	Modifications increase expression per vector genome	135

Supplemental Figure 4	High-performance liquid chromatography (HPLC) chromatograms from peripheral blood lysates obtained from representative mice of each experimental arm	137
Supplemental Figure 5	Peripheral blood and bone marrow engraftment levels	139
<b>Chapter 3</b>		
Figure 1	Overview of LV-MPRA library design and experimental workflow	177
Figure 2	Sequence intrinsic enhancer map spanning the $\beta$ -globin Locus Control Region	180
Figure 3	LV-MPRA guided therapeutic vector design and characterization	182
Figure 4	<i>In vivo</i> analysis of peripheral blood from “Townes” mouse model of SCD	185
Supplemental Figure 1	~100 bp fragment of LCR DNase HS2 element enhances expression from minimal $\beta$ -globin promoter	187
Supplemental Figure 2	Sequence intrinsic enhancer map spanning the $\beta$ -globin Locus Control Region aligns with ENCODE	189
Supplemental Figure 3	Determination of optimal enhancer orientation and arrangement.	191
Supplemental Figure 4	Sequences within the 95th percentile of lowest expression do not enhance expression	193
Supplemental Figure 5	<i>In vitro</i> characterization of vectors designed using LV-MPRA enhancer map	195

Supplemental Figure 6 Analyses of peripheral blood and bone marrow from 197  
“Townes” mouse model of SCD

## Acknowledgements

The work for this dissertation was performed under the direction of Dr. Donald B. Kohn

**Chapter 1:** I would like to thank David Gray and Anastasia Lomova for co-authoring chapter one, which was adapted from Morgan, Richard A., et al. "Hematopoietic stem cell gene therapy: progress and lessons learned." *Cell stem cell* 21.5 (2017): 574-590. DOI: <https://doi.org/10.1016/j.stem.2017.10.010>. Training grants provided support to R.M. (NIH/NHLBI F31 HL134313; Paul Calabresi Predoctoral Fellowship (PhRMA Foundation); UCLA MSTP T32 GM00804) to complete this work.

**Chapter 2:** I would like to thank Mildred J. Unti, Bamidele Aleshe, Devin Brown, Kyle Osborne, Colin Koziol, Oliver B. Smith, Rachel O'Brien, Curtis Tam, Beatriz Campo-Fernandez, Eric Miyahira, and Shantha Senadheera for assistance in designing research materials, planning and executing experiments, and analyzing data. I would like to thank Donald B. Kohn and Roger P. Hollis for assistance in interpreting research results and offering advisement, providing financial and administrative support, and co-authoring the paper. This work was supported by a Sponsored Research Agreement from Biomarin Pharmaceutical, Inc. I would also like to thank the Flow Cytometry Core and the DNA Sequencing Core of the UCLA Eli & Edythe Broad Center of Regenerative Medicine and Stem Cell Research and the Jonsson Comprehensive Cancer Center were used to support studies. Training grants provided support to R.A.M. (NIH/NHLBI

F31 HL134313; Paul Calabresi Predoctoral Fellowship (PhRMA Foundation); UCLA MSTP T32 GM00804). This work has been submitted for peer review.

**Chapter 3:** I would like to thank Feiyang Ma, Mildred J. Unti, Devin Brown, Paul George Ayoub, Curtis Tam, Lindsay Lathrop, Bamidele Aleshe, Ryo Kurita, Yukio Nakamura, Shantha Senedheera, and Matteo Pellegrini for assistance in designing research materials, planning and executing experiments, and analyzing data. I would like to thank Donald B. Kohn and Roger P. Hollis for assistance in interpreting research results and offering advisement, providing financial and administrative support, and co-authoring the paper. This work was supported by a Sponsored Research Agreement from Biomarin Pharmaceutical, Inc. I would also like to thank the Flow Cytometry Core and the DNA Sequencing Core of the UCLA Eli & Edythe Broad Center of Regenerative Medicine and Stem Cell Research and the Jonsson Comprehensive Cancer Center were used to support studies. Additional support was offered by National Center for Advancing Translational Sciences UCLA CTSI Grant UL1TR001881. We would also like to thank the Technology Center for Genomics & Bioinformatics Sequencing Core, Department of Pathology & Laboratory Medicine in the David Geffen School of Medicine at UCLA, for offering technical support. This work was supported by training grants provided to support R.M. (NIH/NHLBI F31 HL134313; Paul Calabresi Predoctoral Fellowship (PhRMA Foundation); UCLA MSTP T32 GM00804). This work has been submitted for peer review.

A special thank you to Dr. Donald B. Kohn for accepting the responsibility of guiding and mentoring me through the PhD. I am eternally grateful for his kindness and the opportunity to have grown under his leadership. I will always owe a debt to him for engraining forever in my mind that “every experiment needs a positive and negative control.”

Completing my PhD would not have been possible without Aaron Marquette, who has taken care of me through some of the most challenging times in my life. Having Aaron around meant that I could get through the toughest of times. Thank you to my aunts, Catherine Gilchrist and Linda Brice for supporting me throughout this entire journey. Lastly, thank you to my mom, Belinda Gilchrist. There was no sacrifice that was too much for her if it would bring me additional opportunity. Thank you for being my first role-model of hard work.

## Richard A. Morgan - Curriculum Vitae

### A. Education

INSTITUTION AND LOCATION	DEGREE	START	END	FIELD OF STUDY
California State University, Northridge, Northridge, CA	BS	08/2004	05/2009	Biology
Johns Hopkins University, Baltimore, MD	MS	08/2009	05/2012	Biotechnology
Charles Drew University of Medicine & Science, Los Angeles, CA	MD/PhD	08/2012	05/2021	Medicine/ Pharm
David Geffen School of Medicine at UCLA, Los Angeles, CA	MD/PhD	08/2012	05/2021	Medicine/ Pharm

### B. Positions and Honors

ACTIVITY	START	END	FIELD	INSTITUTION	SUPERVISOR
Undergraduate Research Assistant	08/07	05/09	Molecular Biology	California State University, Northridge	Mary-Pat Stien, PhD
Research Technician	04/10	05/12	Cancer Genetics	Johns Hopkins University	Christine Iacobuzio-Donahue, MD, PhD
Executive Director; Hopkins Biotech Network	07/10	07/11	Health Policy	Johns Hopkins University	Lynn J. Langer, PhD
Student Coordinator; Pathology Interest Group	01/13	05/14	Pathology	University of California, Los Angeles	Dr. Elena Stark, MD, PhD
Rotating Scholar	05/13	08/13	Cell Biology	University of California, Los Angeles	Amy Rowat, PhD
Graduate Research Assistant	06/2014	2019	Gene Therapy	University of California, Los Angeles	Donald Kohn, MD

ACTIVITY	START	END	FIELD	INSTITUTION	SUPERVISOR
Graduate Student Mentor	06/2015	2019	Counseling	University of California, Los Angeles	Tama Hasson, PhD

### **Academic Honors**

2019 Excellence in Research Award; American Society of Gene and Cell Therapy

2017 Paul Calabresi Scholarship Awardee; PhRMA Foundation

2016 Franklin C. McLean Scholarship Awardee; National Medical Fellowships

2016 Ruth L. Kirschstein National Research Service Award (F31)

2016 Dr. Ursula Mandel Scholarship Awardee; University California Los Angeles

2016 Academic Scholarship Awardee; Philip and Aida Siff Educational Foundation

2015 Victor Grifols Roura Scholarship Awardee; National Medical Fellowships

2014 Alma Wells Givens Scholarship Awardee; National Medical Association

2013 American Society of Clinical Oncology Fellowship Awardee

2013 Jesse B Barber Scholarship Awardee; National Medical Association

2013 Graduate Education Scholarship Awardee; Charles Drew University

2012 Business Plan Competition Finalist; Johns Hopkins University

2009 Medal of Leadership and Service Awardee; Associated Students at California State University Northridge

### **Complete List of Published Work in My Bibliography:**

<http://www.ncbi.nlm.nih.gov/myncbi/richard.morgan.1/bibliography/41605798/public/?sort=date&direction=ascending>

**Chapter 1: Hematopoietic Stem Cell Gene Therapy –Progress and Lessons  
Learned**

## **Abstract**

The use of allogeneic hematopoietic stem cells (HSCs) to treat genetic blood cell diseases has become a clinical standard but is limited by availability of suitable matched donors and potential immunologic complications. Gene therapy using autologous HSCs should avoid these limitations and thus may be safer. Progressive improvements in techniques for genetic correction of HSCs, by either vector gene addition or gene editing, are facilitating successful treatments for an increasing number of diseases. We highlight the progress, successes, and remaining challenges toward development of HSC gene therapies and discuss lessons they provide for development of future clinical stem cell therapies.

## **Introduction**

Most inherited blood cell diseases, such as primary immune deficiencies, hemoglobinopathies, storage and metabolic disorders, congenital cytopenias and stem cell defects, can be treated by transplantation of allogeneic hematopoietic stem cells (HSCs) (**Table 1**) (Boelens et al., 2013; Walters, 2015). The transplanted genetically normal HSCs can serve as an ongoing source of blood cells of all lineages, eliminating these disorders from a single treatment with benefits lasting life-long.

While there are generally high rates of success when an HLA-identical sibling donor is available, the outcomes of hematopoietic stem cell transplantation (HSCT) are usually not as successful with less well-matched allogeneic donors (either haplo-identical family members or unrelated donors) (Boelens et al., 2013; Walters, 2015). Reduced HLA matching between recipient and donor increases the risks of graft rejection and graft

versus host disease (GVHD). Rejection of an HSC graft generally leaves the patient in a perilous position, with an urgent need to restore hematopoiesis to prevent complications from prolonged pancytopenia (anemia, infection, bleeding). The primary donor may not be available (e.g. cord blood units are not linked to their source) and a suitably matched second donor may not be identified. GVHD is a major cause of transplant morbidity and even mortality, and can impose a chronic rheumatologic-like inflammatory/fibrotic disease, with need for persistent immune suppression and the attendant risks of infection and toxicities (Cooke et al., 2017). Immediately before and after the allogeneic transplant, high levels of immune suppression are necessary to reduce immunological risks but these treatments also add to morbidity. There has been ongoing progress with methods to reduce GVHD in allogeneic HSCT, including improved graft engineering by removal of selective T cell populations (TCR  $\alpha/\beta$  depletion, naïve T cell depletion), and by use of post-transplant cyclophosphamide (Fuchs, 2015; Muccio et al., 2016). Nonetheless, immune complications and lack of suitable matched donors present significant clinical barriers to successful application of allogeneic HSCT for a wider range of disorders.

Autologous HSCT in which the patient's HSCs are gene-modified should offer complete avoidance of the major immunological complications of allogeneic HSCT, which may contribute to better outcomes for patients with genetic blood cell disorders. For specific disorders, expression of the gene introduced into HSCs is needed in cells of one or more hematopoietic lineages (e.g. red blood cells, neutrophils, lymphocytes) (**Figure 1**). The lack of immunogenicity with autologous cells allows the use of reduced intensity of the pre-transplant conditioning to make space in the marrow niche to facilitate HSC engraftment, compared to what is required for effective allogeneic HSCT (**Figure 2**).

Current approaches to autologous transplant/gene therapy using lentiviral vectors (LVs) have produced clinical benefits similar to those from allogeneic transplant for several disorders. (Aiuti et al., 2013; Biffi et al., 2013; Cartier and Aubourg, 2010; De Ravin et al., 2016b). In multiple clinical trials (**Table 1**), this approach has consistently achieved quite stable frequencies of gene-corrected blood cells of all lineages, indicating engraftment, long-term persistence and ongoing generative capacity of gene-modified HSCs, with no significant diminution observed over time in human subjects (Cartier et al., 2012; Enssle et al., 2010).

Recent developments in gene editing have led to investigations toward its application for *ex vivo* gene correction in HSCs, which may have advantages compared to integrating viral vector-mediated gene addition (Carroll, 2016; Wright et al., 2016). This review will present the primary approach that is currently being used for gene modification of HSCs for clinical applications and gene addition using integrating viral vectors, as well as discuss the current status of gene editing in human HSCs for autologous transplantation. Lessons learned from advancing HSC therapies to the clinic may help inform the development of other stem cell therapies.

## **HSCs for Gene Therapy**

HSCs are long-lived and multipotent, so gene correction in HSCs should lead to persistent gene correction among the different lineages (Kondo et al., 2003). The hematopoietic system is an ideal target for gene therapy because of the ease with which HSCs can be accessed for *ex vivo* gene manipulation, effective gene-modification, and re-administration as an intravenous infusion

HSCs are traditionally harvested from bone marrow derived from the iliac crests under general anesthesia. Multiple aspirations are performed with the goal of collecting 10-20 ml of bone marrow per kilogram of recipient body weight. Alternatively, HSCs can be obtained as cytokine (e.g. G-CSF)-mobilized peripheral blood stem cells (PBSC) collected by leukopheresis. Hematopoietic growth factors, including GM-CSF and G-CSF, or CXCR4 inhibitors have been shown to increase the numbers of circulating hematopoietic stem and progenitor cells (HSPC) by 30-1000 fold (Brave et al., 2010). PBSCs are now the predominant clinical HSC source used for allogeneic and autologous transplants to routinely and successfully treat multiple blood cell disorders using current techniques.

However, the use of HSCs for gene therapy presents several challenges. HSCs CD34+, CD38-, CD45RA-, CD90+, CD49f+ (Notta et al., 2011), purification to high levels at clinical scale may entail significant losses of cells and impair their stem cell capacity. In current clinical practice for gene therapy, the HSCs from the clinical source (bone marrow or mobilized peripheral blood stem cells) are enriched, rather than purified, usually by isolating the CD34+ fraction using immunomagnetic separation. The CD34+ population (~1% of cells in adult bone marrow) contains most long-term engrafting multipotent HSCs, but also far more numerous short-term progenitor cells. CD34 selection enables ~30-50-fold enrichment of HSCs, removing the majority of highly numerous mature blood cells and enriching the HSC targets to culture for *ex vivo* gene modification. The dosages of CD34-selected cells typically used for transplantation range from 2 to 20 million/kg, necessitating efficient processing of relatively large numbers of cells.

Because they will divide many times, any gene modification of HSCs needs to be permanent and heritable to be passed on to all successive generations of progeny cells. Currently this necessitates making changes in the genome, either by covalent gene addition with an integrating vector or direct genome editing. The critical technical challenge for successful HSC gene therapy is performing sufficient gene engineering of the autologous HSCs to provide a therapeutic level of permanent genetic correction without impairing their stem cell capacity or causing adverse effects.

Thresholds for sufficiency can be based on observations from cases where patients, allo-transplanted for these disorders, develop mixed chimerism with only a sub-fraction of the hematopoiesis coming from donor cells. Clinical improvement has been reported with donor chimerism as low as 10-30% for sickle cell disease, thalassemia, SCID, and other PIDs, making this level a reasonable target for engrafted, gene-corrected HSCs (Chaudhury et al., 2017; Hsieh et al., 2011).

### **Vector Choice and Design**

An attractive property of retroviruses is their ability to convert their RNA genome into proviral DNA through reverse transcription and integration into the DNA of the host cell's genome in a quasi-random fashion. This integrating property of retroviruses allows the transmission of therapeutic information to all progeny of a transduced HSC. The initial retroviral systems used were derived from Murine Leukemia Viruses (MLV) a class of simple gammaretroviruses (gRV) that were well-known from studies of their oncogenic properties. Transduction of human HSCs with gRV vectors has remained challenging due to the quiescent nature of HSCs that typically cycle infrequently during steady state

hematopoiesis (Cheshier et al., 1999; Passegué et al., 2005; Pietras et al., 2011). MLV requires the breakdown of the nuclear envelope and cellular progression through mitosis to stably integrate into host cell genome as the virus lacks active nuclear localization elements (Lewis and Emerman, 1994; D. G. Miller et al., 1990). Therefore, HSCs need to be cultured for several days with multiple cytokines to induce cycling for retroviral transduction, and this may lead to loss of stem cell capacity.

Lentiviral vectors (LVs) have subsequently become the vector platform of choice because they do not require the cells to undergo mitosis for the breakdown of the nuclear membrane to efficiently integrate their proviral DNA into host cells. Rather, LVs transit through the nuclear pores by recruitment of host cell proteins. The most widely used design of LV system used for transduction of HSCs was first developed by Naldini *et al* in 1996 and was subsequently shown to efficiently transduce HSCs by Miyoshi *et al* and Case *et al*, among others (Case et al., 1999; Miyoshi et al., 1999; Naldini et al., 1996). While clinically-effective methods have been developed for gene introduction to human HSCs using retroviral and lentiviral vectors, HSCs are relatively resistant to transduction, requiring the use of high multiplicities of infection of vector (e.g. 10-100, based on titers measured on permissive cells) to effectively modify the majority of HSCs.

### **Safety issues**

The first clinical trial utilizing gene therapy to modify autologous HSCs with curative intent began in 1992. The goal of this first trial conducted by Bordignon *et al* was to correct severe combined immune deficiency (SCID) syndrome caused by deficiency in adenosine deaminase (ADA) (Bordignon et al., 1995). SCID patients experience severe, recurrent

and persistent infections resulting from immunodeficiency and, prior to the availability of HSCT options, the disease was lethal. This first clinical trial employing autologous gene-corrected HSCs to correct ADA-SCID utilized a vector derived from MLV to introduce an ADA cDNA into HSCs isolated from afflicted patients (Aiuti and Roncarolo, 2009; Gaspar et al., 2014). These investigators and other groups in the U.K. and the U.S. have gone on to treat more than 45 ADA SCID patients with gRV vectors, with good immune recovery in most and no complications from the vectors (Candotti et al., 2012; Gaspar et al., 2011; Shaw et al., 2017).

Additional trials using MLV-based gRV vectors, such as those conducted by groups in France and the U.K. for SCIDX1 (X-linked SCID) to correct interleukin 2 common gamma chain (IL2Rg) deficiency, demonstrated both the utility and the limitations of gRV vectors (Hacein-Bey-Abina et al., 2002). Although curative in the majority of patients, five (of 20 total) patients developed T cell acute lymphoblastic leukemia (T-ALL), two to six years post treatment, as a result of the action of vectors that had integrated near proto-oncogenes (Hacein-Bey-Abina et al., 2008; Howe et al., 2008). Similar leukoproliferative complications were seen in other clinical trials using gRV vector for X-linked Chronic Granulomatous Disease (X-CGD) (Ott et al., 2006), and Wiskott-Aldrich Syndrome (WAS) (Boztug et al., 2010). The occurrence of T-ALL or myeloid malignancies in subsets of patients from each of these clinical trials was a result of the LTR driven gRV vector landing upstream of proto-oncogenes and ectopically activating their expression (Hacein-Bey-Abina et al., 2008; Howe et al., 2008). It is now known that gRVs tend to land near transcriptional start sites of genes, CpG islands, and DNase 1 hypersensitive sites (which tend to be transcriptionally active) (De Rijck et al., 2013; Derse

et al., 2007; Emery et al., 2009; Mitchell et al., 2004; X. Wu, 2003). The LTRs of these gRV act as strong enhancers that recruit a number of transcription factors capable of overriding innate cellular transcriptional control of neighboring genes, promoting leukemogenesis (Modlich et al., 2009).

To address MLV's propensity to induce enhancer-mediated insertional mutagenesis, Gilboa *et al* developed the first self-inactivating (SIN) vector by introducing a deletion within the 3' U3 that abolishes enhancer activity. During reverse transcription, the deleted 3' U3 is copied to both ends of provirus DNA and deprives the provirus of LTR mediated enhancer and promoter activities (Yu et al., 1986). Instead, internal promoters can be introduced to drive transgene expression with higher regulated/tissue specific expression. This alteration provides SIN vectors with increased safety by reducing cellular gene activation when in proximity to neighboring promoters (Nienhuis et al., 2006). SIN gRV vectors have been used safely for subsequent studies of gene therapy for SCIDX1, indicating that this modification did achieve its goal of greatly reducing risks of insertional transformation (Hacein-Bey-Abina et al., 2014).

These studies informed the development of LV as vehicles for gene delivery. Self-inactivating deletions were introduced into the viral LTRs of LV and all sequences encoding proteins supporting HIV virulence were deleted from the provirus and all packaging constructs to create second (-Vpr, -Vif, -Vpu, -Nef) (Zufferey et al., 1997) and third (also -Tat) generation LV vector systems (Dull et al., 1998). The VSV-G glycoprotein is most commonly used to pseudotype lentiviral vectors, although other envelope proteins have shown some favorable properties (Girard-Gagnepain et al., 2014). Lentiviral vectors can transduce non-dividing cells via several mechanisms they have for nuclear import of

their viral cores (Matreyek and Engelman, 2013). They also have somewhat larger carrying capacity than gRV vectors (6-9 kb) and are generally more robust for transducing human cells. They have mostly become the vector of choice for stable gene addition to human HSCs.

### **Methods to Optimize Vector Delivery**

A major limitation of gRV vectors is their inability to transduce non-dividing cells efficiently. Addition to culture dishes of a specific adhesion domain of fibronectin in a recombinant protein, CH-206, which recruits virus particles to HSCs (resulting in a higher MOI at the interface between virus and cell) was found to significantly increase transduction efficiency (Hananberg et al., 1996). *Ex-vivo* culture conditions were also found to influence HSC proliferation and transduction efficiency (Barrette et al., 2000; Sutton et al., 1999). Addition of recombinant human hematopoietic growth factors (typically ckit ligand, Flt-T ligand, thrombopoietin, and Interleukin-3) during transduction resulted in activation of CD34+ HSCs and therefore higher transduction rates. The combination of fibronectin and optimal *ex-vivo* culture conditions greatly improved transduction with proven success in the clinic, as discussed below (Millington et al., 2009).

Another method that has been used to enhance transduction is to alter the cell target specificity of a viral vector by exchanging the innate envelope protein for one derived from an alternative virus allows researchers to alter the tropism of resultant vector particle (called pseudotyping). Typically, pseudotype is chosen based on expression level and exclusivity of the envelope protein's cognate receptor (higher receptor levels equal greater gene transfer levels). HSCs can be transduced with RD114, GALV, BaEV and

VSV-G pseudotype viruses, among others. Other, more specific pseudotyping strategies are being developed that include the use of diverse viral envelopes and even fusions proteins with antibodies or cytokines to target specific cell types (Gennari et al., 2009; Verhoeven et al., 2005).

### **Advances in Vector Design**

Surprisingly, there have been no significant improvements to the basic design of lentiviral vectors since the so-called third generation vectors were introduced almost two decades ago (Dull et al., 1998; Zufferey et al., 1998). They were designed based on a decade of experience with gammaretroviral vector design and production and have met all the safety expectations, with no report of emergence of replication-competent lentivirus during packaging in research or clinical manufacturing.

Lentiviral vectors with relatively small and simple gene cassettes (e.g. human phosphoglycerate kinase gene or elongation alpha-1 gene minimal promoters and a cDNA) are readily produced to titers sufficiently high for effective gene modification of human HSCs at clinical-scale. However, low vector titer remains a significant problem with some LVs, especially those tasked to carry larger transgene cassettes, such as the human beta-globin gene. Several studies have shown that increasing viral RNA genome length negatively affects both titer and transduction efficiency (Cant Barrett et al., 2016; Kumar et al., 2004). Reduction of viral RNA length through removal of non-essential sequences is a viable strategy for improving LV titer for large transgene cassettes. Additionally, codon optimization may be used to improve titer by depleting secondary

structures detrimental to mRNA stability (with the added benefit of improving transgene expression levels) (Moreno-Carranza et al., 2008). Other strategies for increasing LV titer include the addition of the Woodchuck Hepatitis Virus Post-transcriptional Regulatory Element (wPRE), which has been shown to increase vector titer through increasing vector genomic RNA stability, export and translation (Hope, 2002; Schambach et al., 2000) or addition of elements known to improve polyadenylation of vector mRNA during packaging, such as the Bovine Growth Hormone Polyadenylation Sequence (Woychik et al., 1984). Recently, Vink et al described a reconfiguration of the basic arrangement of the key cis-regulatory elements of HIV included in the vector backbone (LTRs, primer binding site, rev-responsive element) to simplify the process of reverse transcription, which may be a limit to transduction with large vectors (Vink et al., 2017). The advantages of this design in clinical applications remain to be tested.

To date, lentiviral vectors have been produced by transient transfection of multiple plasmids (vector, packaging proteins, envelope), which is cumbersome and challenging to scale-up to levels that would be needed for commercial production. There have been extensive and long-standing efforts to develop stable lentiviral vector packaging lines, similar to those routinely produced for gammaretroviral vectors, which could provide a master cell bank for much simpler production of vector lots. It has proven to be challenging to repress expression of the VSV-G glycoprotein (which remains the most effective pseudotype for lentiviral vectors in most cases) and some of the HIV-1 genes which may be cytotoxic to the packaging cells for cell passaging, but then rapidly and robustly inducing expression of these genes for vector production. One stable packaging cell line has been used to produce lentiviral vector for clinical trials. (De Ravin et al., 2016b; Throm

et al., 2009) . It was made by successively transfecting plasmids encoding each gene cassette needed to make the virus proteins (HIV gag/pol, VSV-G) under tight, inducible expression control, and then concatamers of the vector plasmid to obtain multiple copies. This led to a cell line that was capable of scale-up to production volumes needed for production of clinical lots without loss of the packaging capacity.

### **Gene Expression from Lentiviral Vectors**

For some gene therapy applications, unregulated constitutive, ubiquitous expression of the transgene is acceptable. For example, adenosine deaminase (ADA) is expressed in all cell types and a broad range of ADA enzyme activity in all blood cell lineages is safe and sufficient to allow immune reconstitution. Thus, the vectors deployed for gene therapy of ADA SCID have used constitutive promoters, such as the MLV LTR or the Elongation Factor 1-Alpha gene promoter (Aiuti et al., 2009; Carbonaro et al., 2014). Other genes may require precise lineage, temporal or physiological-responsive expression patterns to be safe and effective. For example, molecules involved in signal transduction (receptors, intracellular signaling molecules, transcription factors) may be expressed in only specific cell types, or under specific physiologic states or in response to specific stimuli (e.g. *BTK*, CD40 ligand, *JAK3*, Stat proteins) (Brown et al., 1998). In some cases, it has been possible to build vectors using transcriptional control elements from endogenous cellular genes to apply regulated expression of transgenes (e.g. Beta-globin transcriptional control elements directing erythroid-specific expression of beta-globin (Sadelain et al., 2000). However, it may not be possible to incorporate into a vector the necessary regulatory sequences to recapitulate endogenous gene expression

patterns for vectors integrated at an array of chromosomal sites in different cells. Insulator elements have been incorporated into some vectors to attempt to mitigate potential silencing of vector expression by heterochromatinization or trans-activation of adjacent cellular genes (Browning and Trobridge, 2016; Emery, 2011). The benefits of insulators remain theoretical and, in fact, silencing and trans-activation has been a problem with current lentiviral vectors. Endogenous gene expression is often controlled by enhancer and promoter interactions that occur over long distances (up to 100kb) away (West and Fraser, 2005). Regulated gene expression can be achieved to variable extents by adding a gene's known enhancer(s) upstream of a minimal promoter within the LV. These enhancers recruit and bind specific sets of transcription factors to cause *cis* activation of the promoter (Spitz and Furlong, 2012). The best studied example is the use of multiple elements from the  $\beta$ -globin gene locus to achieve erythroid-specific expression for the treatment of hemoglobinopathies (Cavazzana-Calvo et al., 2010; Sadelain et al., 2000). However, it is sometimes difficult to fit all of the necessary *cis*-acting genetic elements required for precise gene expression within the size limit of vector genomes. In some cases, the function of enhancers can be antagonized by repressive chromatin structure at some LV integration sites (Hofmann et al., 2006; Yao et al., 2004). Chromatin domain insulators can be added into the LV's LTRs to overcome these positional effects and may also reduce risks from internal enhancers of vectors from affecting neighboring genes (Ramezani et al., 2008).

## Lentiviral Vectors in Clinical Trials

Clinical trials using LVs began in the mid 2000's (Cartier et al., 2009; Cavazzana-Calvo et al., 2010) and LV have now been used safely and effectively in multiple studies for almost a dozen disorders using hematopoietic stem cells (Table 1) (and in many of the T cell-based immunotherapies with Chimeric Antigen Receptors and T Cell Receptor genes (Morgan and Boyerinas, 2016). In most trials, gene delivery to HSCs was at sufficiently high levels to produce clinically beneficial levels of gene-modified HSCs and relevant mature hematopoietic cells for the treatment of disease. The absence of GVHD and the reduced amounts of conditioning chemotherapy needed for engraftment of the autologous HSCs has allowed the predicted improved safety profiles. Analyses of LV integration site in the blood cells of subjects in the different clinical trials have shown a remarkably consistent pattern, with no predilection for insertion near proto-oncogenes and no clinically significant clonal expansions (Biasco et al., 2016; Biffi et al., 2011). LV gene therapy is being developed for several other disease indications (**Table 1**), including additional primary immune deficiencies, storage and metabolic diseases, and stem cell defects such as Fanconi's Anemia. Limitations to wider applications for more clinical indications are now less biological and more logistic, as each disorder requires its specific vector and the entire pre-clinical drug development pathway. Funding by research grants to perform vector development for additional indications may become more difficult to obtain, as the scientific novelty is diminishing; while funding from pharmaceutical companies and venture capitalists relies on the expectations of financial returns, which are unknown for these relatively rare orphan disorders. Issues related to commercial marketing and reimbursement for these cell and gene therapies are complex (Brennan

and Wilson, 2014; Orkin and Reilly, 2016). Additionally, the capacity to produce the large volumes of clinical-grade LVs for these studies (and also for the larger studies of CAR T cells), is limited and may slow progress. Nonetheless, gene therapy using HSCs is continuing to advance and provide effective and safe therapies for a growing list of disorders.

### **X-linked Adrenoleukodystrophy (X-ALD) HSC Gene Therapy**

A recent report described efficacious HSC gene therapy for X-linked Adrenoleukodystrophy (X-ALD), a progressive neurodegenerative disorder with onset in boys during the first 1-2 decades (Eichler et al., 2017). The major premise for the approach is that the gene-corrected engrafted HSC will produce cells that become CNS microglia and provide essential enzyme activity that can rescue very long chain fatty acid catabolism and prevent demyelination. Graft versus host disease, which seems to accelerate progression of X-ALD (Peters et al., 2004), is avoided with the use of autologous cells. The findings of beneficial stabilization of neurologic status now extend and expand upon those of the initial report of two X-ALD patients treated using a lentiviral vector into CD34+ PBSC\*<sup>i</sup> (Cartier et al., 2009). The larger cohort studied by Eichler et al. (n=17) showed a high rate of response, with only one patient having progressive neurologic deterioration. HSC gene therapy effectively arrested disease progression in

---

Also the first clinical trial to use a lentiviral vector in human CD34+ cell transplants.

88%, either without symptoms or with only initial progression to early neurological impairments followed by stabilization of neuro-imaging studies and clinical function.

This was the first commercially-sponsored clinical trial of HSC gene therapy (bluebird bio), with essentially all prior trials done at academic medical centers as research investigations. It was performed as a Phase II/II trial intended to obtain data to support applications for regulatory approval to market this stem cell gene therapy. Centralized manufacturing of the stem cell product was performed under full GMP conditions. Stem cell mobilization and leukopheresis was performed at several academic clinical sites enrolling subjects and shipped to a contract manufacturing organization for stem cell enrichment and transduction, yielding a cryopreserved product, which was shipped back to the clinical site following completion of product release testing. They produced consistently high quality cell products with good cell dosages ( $6-19 \times 10^6$  CD34+ cells/kg) and gene transduction levels (0.5-2.5 vector copies per cell). Patients were treated within 2-3 months from enrollment, which is an excellent time-frame to produce and certify a gene-modified stem cell product and perform all the clinical evaluations prior to a HSC transplant. The high rate of successful outcomes in terms of halting neurologic progression meet or exceed those of the clinical alternative of unrelated or haplo-identical donor transplants, strongly supporting the use of autologous gene therapy for this disorder (and by extension to many other storage diseases). Bluebird bio is pursuing regulatory drug approval for this combined cell and gene therapy product, which would represent one of the first approved gene therapies (bluebird bio, Inc., 2017).

## Gene Editing

Viral mediated gene transfer has well established benefits and has demonstrated clinical efficiency, as described above. However, risks and drawbacks of these methods remain, such as insertional oncogenesis and modified transgene expression pattern (Cesana et al., 2014; Ng et al., 2010; Zhou et al., 2016). Targeted gene editing allows for site-specific genome modification, and thereby eliminates the risks posed by randomly inserting genes. Another advantage of targeted gene editing over viral vectors is the ability to retain endogenous control of gene expression (Barzel et al., 2014; H. Li et al., 2011). Targeted gene editing can be achieved by employing site-specific endonucleases to induce a double-stranded break (**DSB**) in the DNA near the mutation site. The recruitment of DNA repair proteins to the site of DNA damage stimulates DNA DSB repair via one of two main pathways: non-homologous end joining (**NHEJ**), or homology-directed repair (**HDR**) (**Figure 3**). NHEJ can be used to achieve gene disruption with site-specific induction of a DSB by a targeted nuclease and NHEJ-mediated introduction of small insertions and deletions (**Indels**). NHEJ is a quick but error-prone pathway (Metzger and Iliakis, 2009); HDR is precise, but is dependent on cell cycle phase (S. Kim et al., 2014). Both of these endogenous DNA repair mechanisms can be harnessed for therapeutic benefit. In mammalian cells, NHEJ is more prevalent than HDR (Chiruvella et al., 2013). For a more detailed review on these DNA damage repair pathways, please see (Chapman et al., 2012; Heyer et al., 2010; Hustedt and Durocher, 2016; Lieber et al., 2003; Shrivastav et al., 2008).

Four main classes of endonucleases have been used for gene editing. The first, homing endonucleases recognize DNA sequences up to 40 bp long. These proteins are

naturally found in six structural families (Jasin, 1996). The LAGLIDADG endonuclease family can be engineered to modify the sequence of DNA they recognize. The process is difficult and time consuming. An alternative being explored to facilitate retargeting HEs has been dubbed “megaTAL” (Boissel et al., 2013). These megaTALs feature a DNA binding domain composed of transcriptional activator like (TAL) effector DNA recognition motifs with the active endonuclease domain of a meganuclease.

Zinc finger nucleases (ZFNs), the second generation of engineered targeted endonucleases, provide a more easily modified system than homing endonucleases (Y. G. Kim et al., 1996). The enzyme functions as a dimer, with each ZFN containing three to five Zinc Finger protein motifs which recognize 3 base pair sequences of DNA and half of the *FokI* endonuclease complex. When a pair of ZFNs with appropriate target sequence binds closely enough for their *FokI* domains to dimerize, they make a DSB. ZFNs have promising function but are limited by the complexity of engineering new pairs.

Transcription activator like effector nucleases (TALENs) function similarly to ZFNs but use a different mechanism to recognize specific regions of DNA (Cermak et al., 2011). Instead of the zinc fingers, TALENs have 15-30 repeats of a 35 amino acid transcription activator like effector (TALE). Each TALE is composed of mostly invariable regions with only two amino acid differences known as repeat variable di-residues (RVDs). A TALE recognizes one base pair determined by which RVD a TALE contains. Adding a number of these TALEs together, fusing them to a *FokI* domain, and administering them in pairs, allows similar DSB formation to ZFNs but with an easier and more modular assembly.

The most recently described targeted endonuclease, Cas9, is a monomeric protein guided by a specific type of RNA, known as a CRISPR guide (Mali et al., 2013a). The

guide RNA (gRNA) contains an 18-21-nucleotide long target sequence attached to a 3' RNA scaffold loop for Cas9 protein binding. The target region must be complementary to a region in the DNA immediately upstream of a 2-5 base pair proto-spacer adjacent motif (PAM) which depends on the species of bacteria from which the Cas protein is derived. The most obvious advantage of CRISPR/Cas9 over the other nucleases is the ease and flexibility of developing guides to target new sites.

### **Gene Editing Strategies**

Depending on the disease being targeted, the type of targeted editing required may fall into one of the three categories: gene disruption, gene correction, or gene insertion (**Figure 3**).

Gene Disruption: In certain cases, knocking out a regulatory element, viral receptor or a pathogenic gene may be sufficient to ameliorate the disease-causing phenotype. High levels of gene disruption may be achieved in hematopoietic stem and progenitor cells because this type of editing does not require a donor template and can be done via the NHEJ pathway. For example, disruption of the *BCL11A* erythroid enhancer (a repressor of fetal globin expression) can increase levels of fetal hemoglobin for the treatment of sickle cell disease and beta-thalassemia (Bauer et al., 2013; Bjurström et al., 2016; Canver et al., 2015; Chang et al., 2017). Alternatively, knockout of the *CCR5* gene in cells from HIV-infected individuals can prevent ongoing infection by the virus (Cradick et al., 2013; Hendel et al., 2015; Holt et al., 2010; L. Li et al., 2013; Mandal et al., 2014; J. C. Miller et al., 2010; Perez et al., 2008; Saydaminova et al., 2015; J. Wang et al., 2015). Trials targeting *BCL11A* are approaching the clinic (Chang et al., 2017) and several early

phase clinical trials have been completed using ZFNs to modify the *CCR5* gene in HIV-infected patient peripheral blood T-cells (Tebas et al., 2014) or HSCs (DiGiusto et al., 2016).

Gene Correction: Diseases that result from a single nucleotide substitution or other small genetic lesions may be corrected by providing a homologous donor with the corrective sequence to serve as a template for DNA repair via the HDR pathway. Applications of this approach include correction of the sickle cell mutation in the beta-globin gene and restoration of beta-globin expression in beta-thalassemia. It is challenging to engage HDR-mediated repair in the quiescent, long term HSC population. In pre-clinical studies, gene correction levels of ~10%, ~25%, and ~35% using TALENs, ZFNs and CRISPR/Cas9, respectively, at the HBB locus were reported *in vitro* (DeWitt et al., 2016; Genovese et al., 2014; Hoban et al., 2015). However once the gene-edited cells were transplanted into immunocompromised NOD-scid-IL2Rg<sup>null</sup> (NSG) mice, the gene correction levels decreased to below 10%. These data suggest that correction was less efficient in the true stem cell population than in more differentiated progenitor cells. Further studies must be performed to improve HDR in HSCs. Recently, a new mechanism of gene correction using a single-stranded donor template was reported, which is thought to act in a RAD-51 independent, and fanconi anemia (FA)-dependent manner (Richardson et al., 2017). Modulation of FA pathway may be beneficial for improving the efficiency of gene correction.

Gene Insertion: In many human genetic diseases, there are a variety of different pathogenic mutations spread across the relevant gene in different patients. In general, the lengths of the gene repair tracts mediated by HDR are relatively short (<40 bp), so

that it may be necessary to develop a panel of nuclease/donor template combinations capable of performing efficient editing to cover an entire gene region (Paquet et al., 2016). If this was accomplished using multiple CRISPR guides targeting different sites along a target gene or the genome, each guide might require validation for levels of activity and specificity, which could be impractical for Good Manufacturing Practice (GMP) production of the cells.

Instead, activity of the whole gene can be restored by the targeted insertion of a corrective cDNA of the relevant gene into the start of the endogenous gene locus. Examples of genes being targeted in HSCs by this method include: *IL2Rg* for X-linked Severe Combined Immune Deficiency (X-SCID), *BTK* for X-Linked Agammaglobulinemia (XLA), *CD40L* for X-Linked Hyper IgM Syndrome (XHIM), and *CYBB* for X-Linked Chronic Granulomatous Disease (XCGD) (Clough et al., 2016; De Ravin et al., 2016a; Genovese et al., 2014; Hubbard et al., 2016; Lombardo et al., 2007). Gene cassettes may be inserted into specific gene loci, such as *CCR5* for anti-HIV strategies to knock-out the viral co-receptor gene and simultaneously insert another anti-HIV gene; or into “safe harbors” such as the AAVS1 site which safely supports sustained transgene expression (De Ravin et al., 2016a; DeKolver et al., 2010; Hockemeyer et al., 2009; Lombardo et al., 2007; Mali et al., 2013b; J. Wang et al., 2015).

The donor template or cassette is flanked by homology arms surrounding the nuclease cut site and generally consists of the gene’s full length cDNA complete with a stop codon and a 3’ untranslated region containing the polyadenylation signal. This cDNA donor template is most often delivered via an adeno-associated virus (AAV) or other non-integrating vector (J. Wang et al., 2015). A major challenge of this approach is to achieve

efficient delivery and integration of these larger donor template. In pre-clinical studies, gene insertion rates of up to 43% at the IL2RG, AAVS1 or CYBB loci were achieved *in vitro*; however, as with gene correction, the levels of gene insertion in HSCs decreased *in vivo* (De Ravin et al., 2016a; Genovese et al., 2014; Schirolli et al., 2015; J. Wang et al., 2015).

Recently, two new methods of gene integration were described: homology-independent targeted integration (HITI) and precise integration into target chromosome (PITCh), which use NHEJ and MMEJ machinery, respectively (Nakade et al., 2014; Sakuma et al., 2015; Suzuki et al., 2016). The advantage of these strategies over HDR-mediated gene integration include being able to target cells outside the S/G2 phases of cell cycle.

## **Quiescence**

HSCs are quiescent and mostly reside in the G0/G1 phase of cell cycle, providing a unique challenge for gene editing, since cell cycle phase is a major factor determining which DNA repair pathway is utilized to repair DSB. NHEJ occurs throughout the cell cycle while HDR is mostly restricted to S/G2 phases, when a sister chromatid is available to serve as a homologous template for repair (Branzei and Foiani, 2008; Pietras et al., 2011). This pattern of DNA repair has caused the rates of targeted gene correction and insertion in primary human HSCs to remain relatively low and rates of gene disruption by indels too high. Methods to increase HDR and decrease NHEJ are being developed. Current gene editing protocols utilize culture with a combination of recombinant hematopoietic growth factors (e.g. *ckit* ligand, FLT3 ligand, thrombopoietin and others) to

induce cell cycling 24-72 hours prior to the delivery of nuclease and donor template. However, even with pre-stimulation the majority of HSPCs are not in S/G2 phases. Cell synchronization agents have been used successfully in 293T cells and the H9 embryonic cell line to temporarily arrest the cells in S/G2 phases of cell cycle during DNA repair to increase HDR (Lin et al., 2014). However, the downstream effects of cell synchronization on the self-renewal and differentiation potential of HSCs is not yet known.

An alternative method being explored to improve the precision of gene insertion and gene correction is by simply reducing NHEJ. One such method reported to reduce NHEJ is the inhibition of Ligase IV, which is involved in the final step in the NHEJ pathway (Maruyama et al., 2015; Srivastava et al., 2012; Van Trung Chu et al., 2015). However, other groups have not been able to achieve a significant decrease in NHEJ using this inhibitor (Gutschner et al., 2016; Pinder et al., 2015; Yang et al., 2016). A potential concern with the inhibition of Ligase IV is that decreasing NHEJ levels in the cells may not result in an increase in HDR if the cells have already committed to the end-joining pathway. The effects of this type of late repair pathway blockade are still unknown, but it may lead to lower correction efficiency or even induce apoptosis.

Rather than preventing NHEJ via its terminal step, another possibility is to control the DNA repair pathway choice more upstream at the decision-making stage. For instance, formation of the BRCA1-PALB2-BRCA2 protein complex is crucial for HDR to occur; however, it is inhibited during G1 (Orthwein et al., 2015). Modulating the interaction between BRCA1-PALB2-BRCA2 in U2OS cells allowed Orthwein and colleagues to initiate HDR in the G1 phase of cell cycle (Orthwein et al., 2015). Whether this approach can be translated to primary human HSPCs remains to be tested. Another possibility is

to decrease nuclease cutting in the G1 phase of the cell cycle by adding to Cas9 a fragment of the Geminin protein that causes Cas9 degradation during G1 when only the NHEJ repair pathway is available (Gutschner et al., 2016).

## **Methods of Delivery**

One recurring challenge of targeted editing is how best to deliver the endonuclease and homologous donor template (if necessary for the particular treatment) to HSCs. Primary human HSCs are notoriously resistant to transfection methods of gene delivery (Van Tendeloo, 2001). Electroporation methods to deliver nucleic acids have improved over time and effectively transfer nucleic acid to the majority of HSCs in a treated sample. However, there is often a mild-moderate degree of toxicity from electroporation and this is significantly worsened by delivery of plasmid DNA, in some cases resulting in up to 60% cell death 24 hours post electroporation (Hendel et al., 2015). The delivery of *in vitro* transcribed mRNA encoding the nuclease and either *in vitro* transcribed or chemically-synthesized short guide RNA is better tolerated. Co-delivery of recombinant Cas9 protein complexed to short guide RNA as ribonucleoprotein (RNP) complexes has also been shown to be effective (S. Kim et al., 2014).

Delivery of homologous donor sequences has been achieved with multiple modalities. Chemically-synthesized oligonucleotides (e.g. 50-200 bp in length) are effective donors for small sequence changes, although they may cause moderate toxicity when introduced by electroporation (Hendel et al., 2015). Integrase-defective lentiviral vectors (IDLV), and the more effective adeno-associated virus (AAV) vectors can efficiently deliver donor

sequences of variable lengths (up to several kb) to HSCs with lower cytotoxicity than oligonucleotides or plasmids (Dever et al., 2016; Hoban et al., 2015; J. Wang et al., 2015).

### **Gene Editing of HSCs for Clinical Applications**

For clinical applications, multiple reagents are thus needed to perform gene editing (nuclease and donor) and each will need to be produced under standardized GMP conditions. Research scale editing is typically done with  $0.2-1 \times 10^6$  CD34+ cells per experimental arm; clinical scale will involve at least 5-10x that many CD34+ cells *per kg*, and thus 50-1,000 times more cells. Although standards for acceptable levels of off-target cutting by a nuclease for clinical editing have not been defined, it is incumbent to investigate their occurrence with the most sensitive and relevant assays that can be practically done as part of pre-clinical toxicology assessments.

### **Challenges to Clinical Application of HSCs Gene Therapy**

#### *HSC harvest and expansion*

While there has been much progress in applications of HSC gene therapy, many challenges remain. The numbers of HSCs that can be obtained from a patient are limited by the yields that can be isolated by bone marrow harvest or mobilization, although the combination of G-CSF and a CXCR4 inhibitor (plerixafor) generally leads to abundant cell collections (Brave et al., 2010). Some specific diseases may limit the numbers of HSCs that can be isolated, such as Fanconi anemia, which results in progressive HSC failure, or osteopetrosis, where the marrow space is progressively reduced by the accumulated bone (Daneshbod-Skibba et al., 1980; Giri et al., 2007).

The cell processing manipulations, including stem cell enrichment or gene modification--particularly when using electroporation--may lead to significant cell losses. Efforts to expand the numbers of true transplantable HSCs have been made, with several small molecules (such as SR-1, UM171, PGE<sub>2</sub>) holding some promise, although no massive HSCs expansion has been achieved (Boitano et al., 2010; Fares et al., 2014; Goessling et al., 2011; Hoggatt et al., 2009; North et al., 2007). The goal of producing transplantable HSCs from pluripotent stem cells is advancing, with direct reprogramming to HSCs from endothelial cells also showing promise (Lis et al., 2017; Sugimura et al., 2017).

### *Gene Transfer*

Gene transfer to HSCs has also advanced to a large degree, with current protocols of hematopoietic growth factor stimulation and transduction with lentiviral vectors reaching therapeutic efficacy for many disorders. Nonetheless, here too, improvements are needed. Human HSCs are relatively resistant to lentiviral vectors, evidenced by the seemingly high multiplicities of infection (M.O.I., the vector/cell ratio) needed to effectively transduce HSCs, compared to the relatively easier transduction of the cell lines typically used to gauge vector titers. The carrying capacity of lentiviral vectors has limitations, with vector titers falling off sharply as the size of the gene cassette increases. In our hands, a lentiviral vector at the small end of the size range (e.g. with a simple cDNA or transgene like GFP and a small promoter at ~4 kb proviral length) has a titer 10-30-fold higher than a vector at the large end of the size range (e.g. with a beta-globin gene cassette with exons, introns, upstream locus control region segments at ~9kb). The lower production titer necessitates a proportionately higher volume of vector preparation to produce a

patient dose, increasing the costs. Additionally, the bigger vectors do not transduce HSCs as well as smaller vectors, even when adjusted to matching MOI. Improved transduction of HSCs with lentiviral vectors using small molecules (proteasome inhibitors, cyclosporine A, rapamycin) has been reported in pre-clinical studies (Petrillo et al., 2015; Santoni de Sio, 2006; C. X. Wang et al., 2014), but the effects have not been clinically validated. And, of course, the semi-random integration of the vectors throughout the genome continue to pose genotoxicity risks, although these are greatly diminished with current generation vectors that lack the strong long terminal repeat enhancers that were the major cause of insertional oncogenesis with the first generation of gRV's.

### *Gene Editing*

Gene editing may avoid many of the problems specific to viral vectors, but it too remains less than ideal in several ways. The various site-specific endonucleases (HE, ZFN, TALEN, or CRISPR) are fairly efficient and targeted gene disruption in HSCs is now in clinical trials for HIV (targeting CCR5, HIV co-receptor) and sickle cell disease (targeting BCL11a, repressor of fetal globin) (Chang et al., 2017; DiGiusto et al., 2016); clinical results have not yet been reported. However, the more elegant goals of targeted gene correction and gene insertion are more complicated to achieve, relying on the HDR pathway to perform the desired edits and thus requiring co-delivery of a homologous donor with the nuclease. Before clinical translation of gene editing, GMP methods of gene editing combining multiple GMP-grade reagents (e.g. CRISPR RNP or mRNA and short-guide RNA; AAV vector homologous donor) will need to be established. Pre-clinical work editing human HSCs using a clinically-scale approach has been reported (De Ravin et al., 2017).

### *Ex vivo processing*

The *ex vivo* processing to enrich HSCs for clinical gene therapy has mostly been limited to CD34+ cell selection. This achieves a moderate (30-50-fold) decrease in total numbers of cells that need to be exposed to vector or gene edited (akin to a lineage-negative {lin-} murine population), but yet retains most of the HSCs. However, the CD34+ cell population is still quite heterogeneous with only a small fraction of cells being the target long-lived HSCs. Thus, a large proportion of the vector or gene editing reagents are wasted modifying the more abundant but short-term progenitor cells. Efforts to further enrich for HSCs using additional markers such as CD38(-), CD90(+), CD133(+), etc, (akin to a murine “LSK” {lin-/Sca1+/ckit+ fraction}) have been reported, but require FACS sorting which may entail long processing times and subject the cells to damaging shear forces (Baum et al., 1992). Newer sorting methodologies based on microfluidics or using serial immuno-affinity bead processes may be beneficial if they provide further enrichment without undue losses of cells (Masiuk et al., 2017; H.-W. Wu et al., 2010).

The cell culture methods used in current clinical trials are relatively standardized, using static culture in gas permeable bags or flasks in serum-free medium supplemented with multiple hematopoietic growth factors. The use of continuous feed bioreactors and/or lower partial pressures of oxygen may provide more optimal conditions for HSC modification and preservation. Small molecules such as PGE<sub>2</sub> may also support HSC survival *ex vivo* improving the level of engraftment of gene-modified HSCs (Hoggatt et al., 2009). Additionally, the current *ex vivo* processing of HSCs is often done in multiple open systems, but new closed systems that continuously contain the cells are being

developed that may allow processing to be done in environments less demanding than the current GMP “clean rooms”.

#### *Universal donor products*

Universal donor cells could largely supplant use of autologous cell products if able to achieve the ideal properties of immunogenicity absence. They can be banked as an off-the-shelf, immediately-ready source of compatible normal cells, including regenerative stem cells. Universal donor cells would have a major advantage in that they could be produced in multi-patient dose lots vs. patient-specific single lots using autologous cell products. It is possible to engineer the cells to have favorable properties, e.g. produce a therapeutic protein such as clotting factors and other serum proteins, lysosomal enzymes, anti-tumor T cell receptor or Chimeric Antigen Receptor or an immunomodulative cytokine or chemokine. Allogeneic HSCT sources may continue to have advantages for HSCT for hematologic malignancies due to their potential graft-versus-leukemia effects, although it should become possible to augment specific immune effector cell products for the positive anti-leukemia effect, but without risks for GVHD.

#### *Pre-transplant cytoreductive conditioning*

Finally, the pre-transplant cytoreductive conditioning used to “make space” for engraftment of the isolated and reinfused HSCs is finally advancing beyond the use of cytotoxic chemotherapy drugs or radiation. These agents are effective at ablating the marrow stem cells (myeloablation), which is necessary for engraftment of gene-modified HSCs and for suppressing the recipient’s immune system (immunoablation), which is necessary to avoid immunologic rejection of the graft. However, they may have severe acute toxicities in multiple organ systems (heart, lungs, liver, kidney, GI) and may produce

infertility or sterility, due to toxicity to germ cells. Monoclonal antibodies to HSC surface proteins (e.g. ckit, CD47, CD45) have been shown in murine models to allow improved engraftment without apparent toxicity (Chhabra et al., 2016; Czechowicz et al., 2007; Palchaudhuri et al., 2016; Xue et al., 2010). These efforts are now being translated to the clinic and may eliminate the need to use toxic preparative regimens to facilitate engraftment.

## **Lessons Learned**

### **It takes a long time**

Lessons learned from the almost 30-year history of developing clinical HSC gene therapy products can inform emerging stem cell-based cellular therapies for myriad other non-hematopoietic diseases such as Duchene's Muscular Dystrophy, Huntington's Disease, Parkinson's Disease, Diabetes Mellitus and others. One clear lesson has been that development of novel therapies takes a long time. Methods for effective gene transfer to HSCs were initially developed in the 1980s and clinical trials started in the 1990's. The initial trials yielded no evidence of efficacy; the first clinical successes were not seen until the 2000's and only in the last decade are therapeutic benefits being conferred consistently for multiple disorders. Several promising HSC gene therapy cell products are advancing through early phase clinical trials (for indications including X-adrenoleukodystrophy, Metachromatic Leukodystrophy, Beta-thalassemia, Sickle Cell Disease, ADA-deficient and X-linked forms of SCID, Wiskott-Aldrich Syndrome, Chronic Granulomatous Disease) and towards licensure for commercial manufacture and sales, with the first (Strimvelis for ADA-deficient SCID) approved by the European Medicines

Agency (European Medicines Agency, 2016). While it may be expected that therapies using other stem cell types will be derived in a shorter time-frame, drug development remains a slow process.

### **Support academic medical centers to develop novel cell therapies**

HSC gene therapies were incubated at academic medical centers in multiple countries often at innovative HSCT programs, not via the traditional pharmaceutical company model of drug development. Universities and other research centers need to have sufficient infrastructure for early phase clinical trial performance and GMP cell processing to achieve similar academic pioneering for other stem cell therapies. Indeed, the centers that have had strong gene therapy programs with the necessary cell processing and regulatory infrastructure have been the leaders in this field (e.g. TIGET, Milan Italy; Hôpital Necker-Enfants Malade, Paris, France; University College London, London, UK; The National Institutes of Health, Bethesda MD; University of California, Los Angeles, Los Angeles CA; Boston Children's Hospital, Boston MA; St. Jude Children's Research Hospital, Memphis TN). Ongoing support of this type of research will be essential to continue the innovation of new therapies.

### **Do it right**

In the early days for the field of gene therapy, the NIH RAC-provided public oversight to review clinical protocols to be performed in academic medical centers was an important forum to discuss the novel potential biohazard issues, as well as scientific and ethical concerns. This may have tempered or even slowed some advancements, but it helped

provide another level of expertise, in addition to that provided by local IRB and FDA (and EMA in Europe), to ensure that trials were based on sound scientific principles, had adequate supporting pre-clinical data on potential efficacy and safety, and were well-designed and monitored. The role of the RAC in overseeing individual trials has decreased, but it still serves its federal advisory role in assessing novel biosafety issues. Other forms of stem cell therapy should also proceed with a base of strong pre-clinical data, careful consideration of the clinical setting and approach, as well as well-controlled cell manufacturing and regulatory oversight, to provide maximum safety for subjects and quality of data derived.

### **Protect subject safety by strict compliance**

Any clinical trial with a novel major intervention, such as cell therapy, and especially with subjects with organ dysfunction caused by their disease, can have unexpected and potentially fatal events, either related to the cell product or not. The death of a volunteer subject in a gene therapy trial in 1999 shocked the field and the effects extended throughout much of academic clinical research (Gelsinger and Shamoo, 2008; Wilson, 2009). The response was to strengthen the quality of clinical trial performance to provide maximum protection to subjects and preserve the integrity of the data they contribute (Wilson, 2009). The field of HSCT began clinical investigations in the 1960's-70's, prior to the establishment of IRB and the other oversight bodies, as best available clinical practice for severe, generally fatal disorders. Since that time, the standards for clinical investigations have been greatly expanded for a much more complex regulatory environment. A typical clinical trial of gene therapy may undergo review by a dozen or

more entities, including IRB, IBC, ISPRC, DSMB, NIH RAC, FDA, and one or more funding agencies. HSCT came to full maturity in the U.S. by the development, initiated by members of the ASBMT academic society, of the Foundation for the Accreditation of Cellular Therapy (FACT), which brought uniformity and high standards for cell processing, clinical operations, data management, regulatory management and other clinical trial activities. Commercial cell processing methods and standards have also been developed supporting several cell products that advanced to relatively late stage of investigation (e.g. neural progenitor cells derived from fetal tissue or pluripotent stem cells) and producing a licensed dendritic cell vaccine (the marketed prostate cancer dendritic cell Sipuleucel-T from Dendreon Corp.). This industry has a strong base to produce high quality cell products, but each new cell product type developed brings unique challenges, including details of the cell processing protocol, the release testing, storage, transport, and therapy administration.

### **Translational research is drug development**

It is important to keep in mind that the goal of translational research is to develop a drug that is effective and safe to achieve licensure, be it a cell and/or gene therapy product such as lentiviral-transduced HSCs, iPSC derived myoblasts, dopaminergic neurons, shRNA, etc. This necessitates basic scientists learning fundamental principles of drug development, such as Good Laboratory Practices, Good Manufacturing Practice and Good Clinical practice. An important tool commonly used in drug development is the Target Product Profile (TPP), which sets goals for attributes like: clinical indication, patient population, administration route and schedule, clinical efficacy targets, potential risks,

drug quality and testing methods. Early drafting of a TPP can guide subsequent studies to keep focused on the drug development goals.

### **Trials should be designed to be informative about the cell product**

While the primary end-points for early phase trials mostly relate to safety, it should be possible to incorporate secondary end-points for efficacy and exploratory end-points for biomarkers that can be assessed for potential suitability as eventual primary end-points for drug approval. In the clinical trials we have done testing new vectors, there has been a primary end-point assessing safety, which is typical for a Phase I study, but also secondary end-points assessing efficacy. For ADA SCID, this has involved safety assessments by documentation of clinical adverse events, as well as ensuring absence of replication-competent viral vector emergence and absence of vector-driven clonal expansion. Efficacy assessments involved measuring expression of ADA enzyme activity in mature blood cells, quantifying engraftment of gene-modified stem cells by measuring vector copy number in cells by quantitative PCR, and performing standard clinical tests of immune function, as well as recording clinical health.

In some instances, especially with orphan diseases, non-traditional pathways of clinical trials may be accepted by regulators, with even small trials used as pivotal for registration, assuming they were done with appropriate design and rigor. The EMA approval of Strimvelis for ADA SCID was based on a single center's Phase I/II clinical data involving 12 patients which served as a pivotal clinical trial.

### **Get the most from pre-clinical studies**

The other major serious complication in the gene therapy field was the development of leukemia in subjects in several primary immune deficiency trials from insertional oncogenesis by retroviral vectors introduced into HSCs (Braun et al., 2014; Hacein-Bey-Abina et al., 2003; Stein et al., 2010). The relatively high frequency of the development of leukemia in some trials (25-75%) was not predicted by pre-clinical studies. However, pre-clinical models may not detect clinical risks that can occur in patients with much larger absolute cell dosages and longer post-treatment time periods. It may be difficult to test cell therapies by the parameters traditionally applied to drug therapies (pharmacokinetics, biodistribution, toxicity), but these aspects can be often assessed using PCR methods to quantify transgenes or cellular markers, such as human genomes in human cells against the background of murine host genomes.

Nonetheless, pre-clinical studies that are performed should be optimized to provide as much relevant information as possible. Principles of Good Laboratory Practices should be applied whenever possible, even at early stages of discovery and prior to formal IND-enabling studies. These include such key elements of GLP as following a detailed pre-defined plan for the studies, statistical plan, data capture forms, with formal data reporting. Again, even during early phases of product development, it important to include toxicology analysis within efficacy studies to obtain initial information that can be used in the design of definitive studies for IND application. Preliminary proof-of-principles studies can also be used to begin to investigate cell dosages, potential toxicities, as well as disease-modifying activity.

## **Advancing clinical cell therapies is challenging**

Clinical cell therapy requires point-to-point control of the manufacturing process and starting materials (e.g. from skin biopsy to delivery of iPSC-derived somatic cell product, whether it is HSCs or other cell product). The GMP process requires highly trained staff and SOPs, materials specification, batch records, personnel training, in addition to the highly-controlled environment and regulated processes. To characterize the cell product for human administration, it is necessary to define release criteria – identity, purity, potency, and safety. A Certificate of Analysis is completed for each batch of cell product and the testing for each critical attribute is required with full documentation. The analytic testing for aspects of cell quality (e.g. cell counts and viability, immunohistochemistry or flow cytometry, PCR, RNA-SEQ, etc.) should be performed using well-characterized assays, which should be made more robust with advancing stages of investigation.

## **Conclusion**

In conclusion, gene therapy using HSCs has progressed over three decades from ineffectiveness to being able to essentially cure several different disorders. The pathway was not linear, but required multiple iterative bench-to-bedside cycles. It is likely that therapies using other stem cells will also have progress and set-backs. But, because the underlying hypotheses for cellular therapies are so convincing, it is highly likely that multiple novel stem cell-based therapies will be developed. The lessons from the field of HSC gene therapy may provide some guidance for investigators pursuing the translational process.

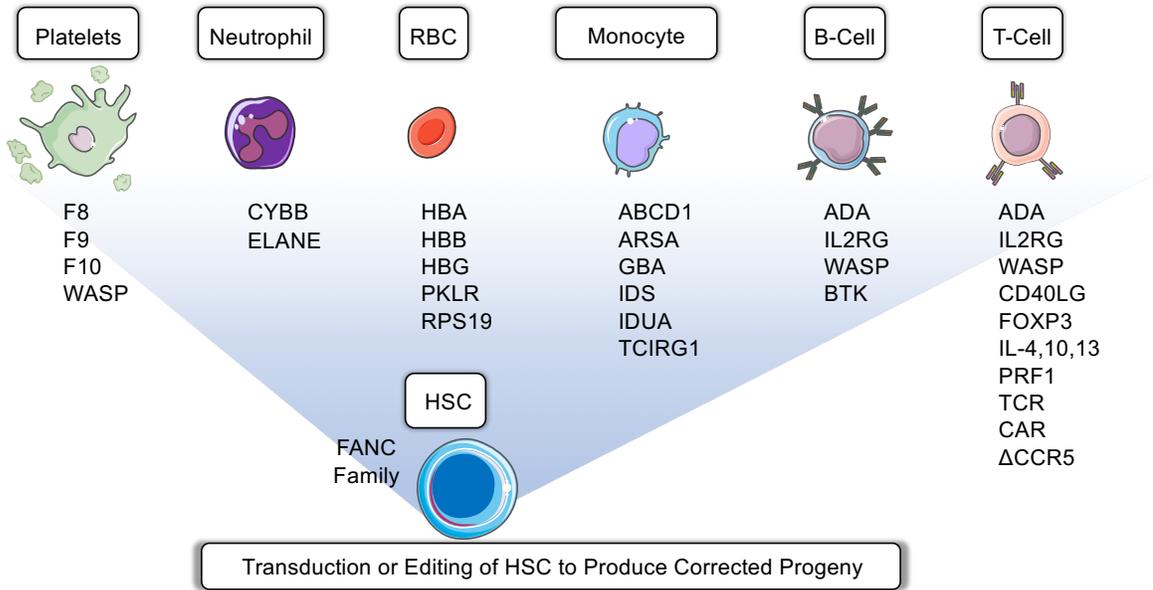
**Tables, Figures and Legends:**

**Table1: Genetic Diseases of Blood Cells and the Transplantation Modalities that Have Been Applied Clinically as Therapies or Are in Pre-clinical Development**

Category of Disease	Specific Conditions	Transplantation Modalities Applied			
		Allogeneic HSCT	Y-Retroviral Gene Therapy	Lentiviral Gene Therapy	Genome Editing
Primary immune deficiencies	ADA-deficient severe combined immune deficiency	+	+	+	n.d
	X-linked severe combined immune deficiency	+	+	+	pre-clinical
	other genetic forms of SCID (Artemis, Rag1/2).	+	n.d	pre-clinical	n.d
	Wiskott-Aldrich Syndrome	+	+	+	pre-clinical
	chronic granulomatous disease	+	+	+	pre-clinical
	leukocyte adhesion deficiency	+	+	pre-clinical	n.d
	hemophagocytic lymphohistiocytosis	+	n.d	pre-clinical	n.d
	X-linked hyper IgM syndrome	+	n.d	n.d	pre-clinical
	X-linked lymphoproliferative disease	+	n.d	n.d	n.d
	X-linked Agammaglobulinemia	Few	n.d	n.d	pre-clinical
	common variable immunodeficiency	heterogeneous genetic etiologies, often unknown; may require editing	heterogeneous genetic etiologies, often unknown; may require editing	heterogeneous genetic etiologies, often unknown; may require editing	heterogeneous genetic etiologies, often unknown; may require editing
	sickle cell disease	+	n.d	+	pre-clinical
	beta-thalassemia	+	n.d	+	pre-clinical
	Storage and metabolic disorders	Gaucher Disease and other lipidoses	+	+	n.d
mucopolysaccharidoses (I-VII)		+	n.d	+	n.d
X-linked Adrenoleukodystrophy		+	n.d	+ recent phase II/III	n.d
metachromatic leukodystrophy		+	n.d	+	n.d
osteopetrosis		+	n.d	pre-clinical	n.d
Congenital cytopenias and stem cell defects	Fanconi anemia	+	+	+	pre-clinical
	Schwachman-Diamond Syndrome	+	n.d	pre-clinical	pre-clinical
	Kostmann's Syndrome	+	n.d	n.d	n.d

**Figure 1: Overview of targets for gene therapy.** Hematopoietic stem cells (HSC) isolated from bone marrow can be modified ex vivo and transferred to recipient to produce functional terminally differentiated cells. Alternatively, differentiated T cells, B cells, or Monocytes can be isolated and modified ex vivo and transferred to the patient. Specific cellular targets and the relevant diseases and genes for gene therapy include the following: **Hematopoietic Stem Cells:** Fanconi Anemia (FANC A-F). **Platelets:** Factor VIII (F8), Hemophilia A; Factor IX (F9), Hemophilia B; Factor X (F10), Factor X deficiency; Wiskott-Aldrich Syndrome Protein (WASP). **Neutrophils:** Cytochrome B-245 Beta Chain (CYBB), chronic granulomatous disease; Elastase Neutrophil Expressed (ELANE), Kostmann's Syndrome. **Erythrocytes:** Hemoglobin Subunit Alpha (HBA), Alpha-Thalassemia; Hemoglobin Subunit Beta (HBB), Beta-Thalassemia and Sickle Cell Disease; Pyruvate Kinase, Liver and RBC (PKLR), Pyruvate Kinase Deficiency; Ribosomal Protein S19 (RPS19), Diamond-Blackfan anemia. **Monocytes:** ATP Binding Cassette Subfamily D Member 1 (ABCD1), X-linked Adrenoleukodystrophy; Arylsulfatase A (ARSA), Metachromatic Leukodystrophy; Glucosylceramidase Beta (GBA), Gaucher disease; Iduronate 2-Sulfatase (IDS), Hunter Syndrome; Iduronidase, Alpha-L (IDUA), Mucopolysaccharidosis type I; T-Cell Immune Regulator 1 (TCIRG1), Osteopetrosis. **B Cells:** Adenosine Deaminase (ADA), Severe Combined Immunodeficiency; Interleukin 2 Receptor Subunit Gamma (IL2RG), X-linked severe combined immunodeficiency; Wiskott-Aldrich Syndrome Protein (WASP); Bruton Tyrosine Kinase (BTK), X-linked agammaglobulinemia. **T Cells:** Adenosine Deaminase (ADA), Severe Combined Immunodeficiency; Interleukin 2 Receptor Subunit Gamma (IL2RG), X-linked severe combined immunodeficiency; Wiskott-Aldrich Syndrome

Protein (WASP); CD40 Ligand (CD40L), X-linked hyper IgM syndrome; Forkhead Box P3 (FOXP3), IPEX Syndrome; Interleukin 4, 10, 13 (IL-4, 10, 13) – Inflammatory disease; Perforin 1 (PRF1), X-linked lymphoproliferative disease; Artificial T cell receptors (TCR), Cancer; Chimeric Antigen Receptor (CAR), Cancer; C-C Motif Chemokine Receptor 5 (CCR5), Human immunodeficiency virus.



**Figure 2. Autologous hematopoietic stem cell transplantation combined with gene addition or editing.**

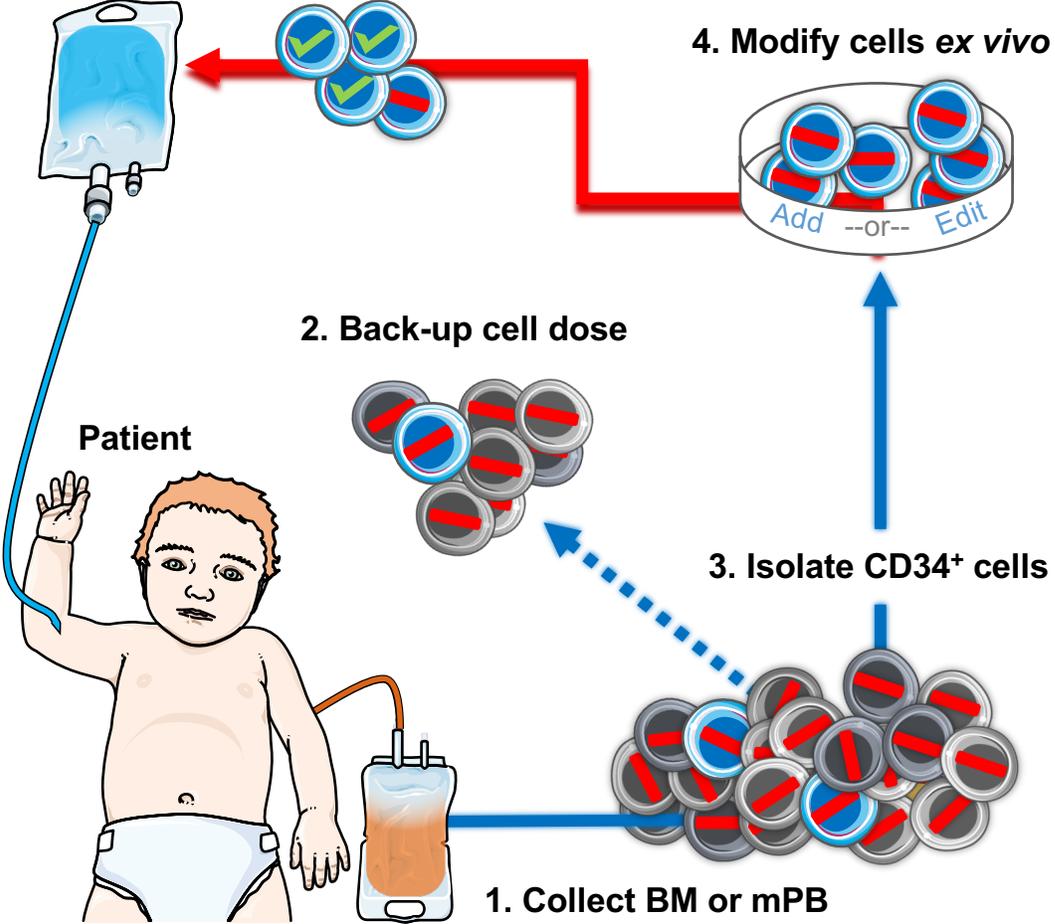
(1) Bone marrow (BM) or mobilized peripheral blood (mPB) cells are collected from the patient (red line represents a disease-causing mutation).

Typically, 15-20ml of BM/Kg is an acceptable harvest target. While collecting HSCs by mobilization and apheresis is less invasive than BM aspiration, infants have small blood volumes making leukapheresis challenging. Failure to harvest adequate cell numbers can prevent therapy. (2) Modification of HSCs may reduce stem cell capacity. A back-up cell dose of non-modified cells is apportioned to restore native hematopoiesis in the event of graft failure. (3) CD34+ cells are isolated in a GMP-compliant, closed system. Purification of HSCs may reduce total cell number as CD34+ HSCs represent less than one percent of total cells. Alternatively, a CD34+/CD38- enrichment strategy may be employed to further purify HSCs and lower the amount vector required for modification. CD34+ cells may be pre-stimulated ex vivo for 1-3 days prior to modification, depending on the protocol. (4) Gene modification of HSCs must be permanent so as to be passed down to all progeny. Cells are modified by either a viral vector to add a gene (typically requires high concentration vector), or targeted nucleases with/without a donor template to disrupt, correct, or insert a gene. After ex vivo modification, the cell product undergoes release testing to assess purity, identity, safety, potency (transduction/editing efficiency), and other characteristics. If the modification strategy requires selection of corrected cells, low cell yield may prevent transplantation. (5) Prior to receiving the cell product, the patient undergoes conditioning to “make space” for engraftment of modified HSCs (green check represents successful modification of a disease-causing gene).

Modified cells may be reinfused fresh or cryopreserved for delivery at a later time. While

high-levels of cytoreductive agents may be toxic, inadequate conditioning may result in poor engraftment.

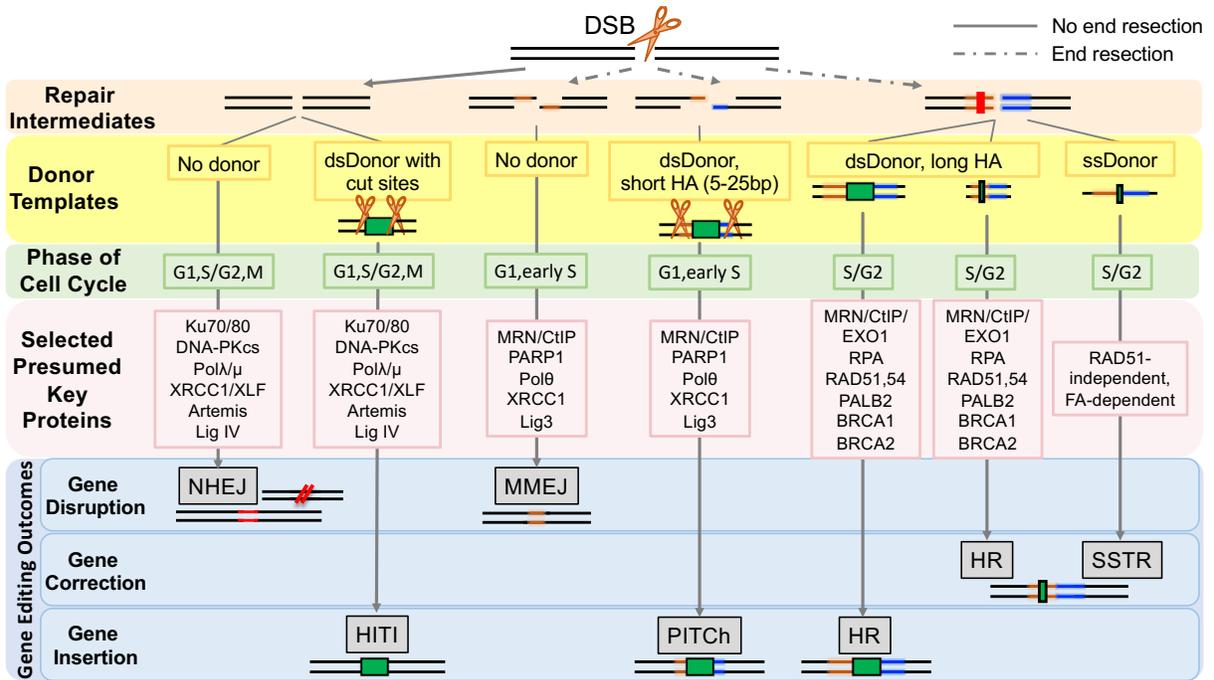
**5. Reinfuse modified cells**



### **Figure 3: Summary of Gene Editing Pathways.**

Double stranded break (DSB) is induced by a targeted nuclease (represented by scissors). DSB ends may or may not be resected (dashed or solid line, respectively),. The ultimate gene editing outcome (light blue boxes on the bottom) depends on several factors: the type of donor template provided (yellow box), the phase of cell cycle (light green box) and the presumed DNA repair proteins available (pink box). Gray boxes indicate the names of the repair mechanisms. It should be noted that the figure illustrates the common pathways described to date, however modification of DNA repair pathways and their utilization for gene editing purposes is an area of active research. (From left to right). A DSB with no end resection and no donor available is likely to result in insertions and deletions (indels) and lead to gene disruption via the non-homologous end joining (NHEJ) pathway. NHEJ may occur in any phase of cell cycle. Exogenously providing a double-stranded donor (dsDonor), which contains nuclease cut sites (scissors) around the gene of interest (green rectangle), may result in homology-independent targeted integration (HITI). The presence of microhomology on opposite strands of DNA around the cut site may result in gene disruption via the microhomology-mediated end joining (MMEJ) pathway. A recently reported method of gene integration, termed precise integration into target chromosome (PITCh), utilizes MMEJ machinery to integrate a gene of interest, which is provided by dsDonor with short homology arms (HA) to the DNA (HA are highlighted in orange and blue). The three pathways on the right are generally only active in S/G2 phases of cell cycle and may be used to correct a single nucleotide mutation in the DNA (represented by a red line). Exogenously providing a dsDonor with long homology arms may lead to either gene integration or

gene correction via homologous recombination (HR) mechanism, depending on the length of the donor template. A new type of repair mechanism for gene correction was recently described, termed single stranded template repair (SSTR). Although resulting in the same outcome as HR-mediated gene correction, SSTR is presumed to utilize the Fanconi Anemia (FA) pathway and be RAD-51 independent.



## References

1. Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M.P., Baricordi, C., Dionisio, F., Calabria, A., Giannelli, S., Castiello, M.C., Bosticardo, M., Evangelio, C., Assanelli, A., Casiraghi, M., Di Nunzio, S., Callegaro, L., Benati, C., Rizzardi, P., Pellin, D., Di Serio, C., Schmidt, M., Kalle, von, C., Gardner, J., Mehta, N., Neduva, V., Dow, D.J., Galy, A., Miniero, R., Finocchi, A., Metin, A., Banerjee, P.P., Orange, J.S., Galimberti, S., Valsecchi, M.G., Biffi, A., Montini, E., Villa, A., Ciceri, F., Roncarolo, M.G., Naldini, L., 2013. Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome. *Science* 341, 1233151–1233151. doi:10.1126/science.1233151
2. Aiuti, A., Cattaneo, F., Galimberti, S., Benninghoff, U., Cassani, B., Callegaro, L., Scaramuzza, S., Andolfi, G., Mirolo, M., Brigida, I., Tabucchi, A., Carlucci, F., Eibl, M., Aker, M., Slavin, S., Al-Mousa, H., Ghonaium, Al, A., Ferster, A., Duppenhaler, A., Notarangelo, L., Wintergerst, U., Buckley, R.H., Bregni, M., Markt, S., Valsecchi, M.G., Rossi, P., Ciceri, F., Miniero, R., Bordignon, C., Roncarolo, M.G., 2009. Gene therapy for immunodeficiency due to adenosine deaminase deficiency. *N Engl J Med* 360, 447–458. doi:10.1056/NEJMoa0805817
3. Aiuti, A., Roncarolo, M.G., 2009. Ten years of gene therapy for primary immune deficiencies. *Hematology* 2009, 682–689. doi:10.1182/asheducation-2009.1.682
4. Barrette, S., Douglas, J.L., Seidel, N.E., Bodine, D.M., 2000. Lentivirus-based vectors transduce mouse hematopoietic stem cells with similar efficiency to

Moloney murine leukemia virus–based vectors. *Stem Cells* 96, 3385–153.

doi:10.1634/stemcells.18-2-152

5. Barzel, A., Paulk, N.K., Shi, Y., Huang, Y., Chu, K., Zhang, F., Valdmann, P.N., Spector, L. P., Porteus, M.H., Gaensler, K.M., Kay, M.A., 2014. Promoterless gene targeting without nucleases ameliorates haemophilia B in mice. *Nature* 517, 360–364. doi:10.1038/nature13864
6. Bauer, D.E., Kamran, S.C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., Shao, Z., Canver, M.C., Smith, E.C., Pinello, L., Sabo, P.J., Vierstra, J., Voit, R.A., Yuan, G.C., Porteus, M.H., Stamatoyannopoulos, J.A., Lettre, G., Orkin, S.H., 2013. An Erythroid Enhancer of BCL11A Subject to Genetic Variation Determines Fetal Hemoglobin Level. *Science* 342, 253–257. doi:10.1126/science.1242088
7. Baum, C.M., Weissman, I.L., Tsukamoto, A.S., Buckle, A.M., Peault, B., 1992. Isolation of a candidate human hematopoietic stem-cell population. *Proceedings of the National Academy of Sciences* 89, 2804–2808. doi:10.1073/pnas.89.7.2804
8. Biasco, L., Pellin, D., Scala, S., Dionisio, F., Basso-Ricci, L., Leonardelli, L., Scaramuzza, S., Baricordi, C., Ferrua, F., Cicalese, M.P., Giannelli, S., Neduva, V., Dow, D.J., Schmidt, M., Kalle, von, C., Roncarolo, M.G., Ciceri, F., Vicard, P., Wit, E., Di Serio, C., Naldini, L., Aiuti, A., 2016. In Vivo Tracking of Human Hematopoiesis Reveals Patterns of Clonal Dynamics during Early and Steady-State Reconstitution Phases. *Cell Stem Cell* 19, 107–119. doi:10.1016/j.stem.2016.04.016
9. Biffi, A., Bartolomae, C.C., Cesana, D., Cartier, N., Aubourg, P., Ranzani, M.,

- Cesani, M., Benedicenti, F., Plati, T., Rubagotti, E., Merella, S., Capotondo, A., Sgualdino, J., Zanetti, G., Kalle, von, C., Schmidt, M., Naldini, L., Montini, E., 2011. Lentiviral vector common integration sites in preclinical models and a clinical trial reflect a benign integration bias and not oncogenic selection. *Blood* 117, 5332–5339. doi:10.1182/blood-2010-09-306761
10. Biffi, A., Montini, E., Lorioli, L., Cesani, M., Fumagalli, F., Plati, T., Baldoli, C., Martino, S., Calabria, A., Canale, S., Benedicenti, F., Vallanti, G., Biasco, L., Leo, S., Kabbara, N., Zanetti, G., Rizzo, W.B., Mehta, N.A.L., Cicalese, M.P., Casiraghi, M., Boelens, J.J., Del Carro, U., Dow, D.J., Schmidt, M., Assanelli, A., Neduva, V., Di Serio, C., Stupka, E., Gardner, J., Kalle, von, C., Bordignon, C., Ciceri, F., Rovelli, A., Roncarolo, M.G., Aiuti, A., Sessa, M., Naldini, L., 2013. Lentiviral Hematopoietic Stem Cell Gene Therapy Benefits Metachromatic Leukodystrophy. *Science* 341, 1233158–1233158. doi:10.1126/science.1233158
11. Bjurström, C.F., Mojadidi, M., Phillips, J., Kuo, C., Lai, S., Lill, G.R., Cooper, A., Kaufman, M., Urbinati, F., Wang, X., Hollis, R.P., Kohn, D.B., 2016. Reactivating Fetal Hemoglobin Expression in Human Adult Erythroblasts Through BCL11A Knockdown Using Targeted Endonucleases. *Molecular Therapy - Nucleic Acids* 5, e351. doi:10.1038/mtna.2016.52
12. Boelens, J.J., Aldenhoven, M., Purtil, D., Ruggeri, A., DeFor, T., Wynn, R., Wraith, E., Cavazzana-Calvo, M., Rovelli, A., Fischer, A., Tolar, J., Prasad, V.K., Escolar, M., Gluckman, E., O'Meara, A., Orchard, P.J., Veys, P., Eapen, M., Kurtzberg, J., Rocha, V., on behalf of all participating centers from Eurocord, Inborn Errors Working Party of European Blood and Marrow Transplant group,

- Duke University, and the Centre for International Blood and Marrow Research, 2013. Outcomes of transplantation using various hematopoietic cell sources in children with Hurler syndrome after myeloablative conditioning. *Blood* 121, 3981–3987. doi:10.1182/blood-2012-09-455238
13. Boissel, S., Jarjour, J., Astrakhan, A., Adey, A., Gouble, A., Duchateau, P., Shendure, J., Stoddard, B.L., Certo, M.T., Baker, D., Scharenberg, A.M., 2013. megaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. *Nucleic Acids Res* 42, 2591–2601. doi:10.1093/nar/gkt1224
14. Boitano, A.E., Wang, J., Romeo, R., Bouchez, L.C., Parker, A.E., Sutton, S.E., Walker, J.R., Flaveny, C.A., Perdew, G.H., Denison, M.S., Schultz, P.G., Cooke, M.P., 2010. Aryl Hydrocarbon Receptor Antagonists Promote the Expansion of Human Hematopoietic Stem Cells. *Science* 329, 1345–1348. doi:10.1126/science.1191536
15. Bordignon, C., Notarangelo, L.D., Nobili, N., Ferrari, G., Casorati, G., Panina, P., Mazzolari, E., Maggioni, D., Rossi, C., Servida, P., Ugazio, A.G., Mavilio, F., 1995. Gene Therapy in Peripheral Blood Lymphocytes and Bone Marrow for ADA- Immunodeficient Patients. *Science* 270, 470–475. doi:10.1126/science.270.5235.470
16. Boztug, K., Schmidt, M., Schwarzer, A., Banerjee, P.P., Díez, I.A., Dewey, R.A., Böhm, M., Nowrouzi, A., Ball, C.R., Glimm, H., Naundorf, S., Kühlcke, K., Blasczyk, R., Kondratenko, I., Maródi, L., Orange, J.S., Kalle, von, C., Klein, C., 2010. Stem-Cell Gene Therapy for the Wiskott–Aldrich Syndrome. *N Engl J Med* 363, 1918–1927. doi:10.1056/NEJMoa1003548

17. Branzei, D., Foiani, M., 2008. Regulation of DNA repair throughout the cell cycle 9, 297–308. doi:10.1038/nrm2351
18. Braun, C.J., Boztug, K., Paruzynski, A., Witzel, M., Schwarzer, A., Rothe, M., Modlich, U., Beier, R., Göhring, G., Steinemann, D., Fronza, R., Ball, C.R., Haemmerle, R., Naundorf, S., Kühlcke, K., Rose, M., Fraser, C., Mathias, L., Ferrari, R., Abboud, M.R., Al-Herz, W., Kondratenko, I., Maródi, L., Glimm, H., Schlegelberger, B., Schambach, A., Albert, M.H., Schmidt, M., Kalle, von, C., Klein, C., 2014. Gene therapy for Wiskott-Aldrich syndrome--long-term efficacy and genotoxicity. *Molecular Therapy* 22, S106. doi:10.1016/S1525-0016(16)35290-X
19. Brave, M., Farrell, A., Ching Lin, S., Ocheltree, T., Pope Miksinski, S., Lee, S.-L., Saber, H., Fourie, J., Tornoe, C., Booth, B., Yuan, W., He, K., Justice, R., Pazdur, R., 2010. FDA Review Summary: Mozobil in Combination with Granulocyte Colony-Stimulating Factor to Mobilize Hematopoietic Stem Cells to the Peripheral Blood for Collection and Subsequent Autologous Transplantation. *Oncology* 78, 282–288. doi:10.1159/000315736
20. Brennan, T.A., Wilson, J.M., 2014. The special case of gene therapy pricing. *Nature Biotechnology* 32, 874–876. doi:10.1038/nbt.3003
21. Brown, M.P., Topham, D.J., Sangster, M.Y., Zhao, J., Flynn, K.J., Surman, S.L., Woodland, D.L., Doherty, P.C., Farr, A.G., Pattengale, P.K., Brenner, M.K., 1998. Thymic lymphoproliferative disease after successful correction of CD40 ligand deficiency by gene transfer in mice. *Nat Med* 4, 1253–1260. doi:10.1038/3233
22. Browning, D., Trobridge, G., 2016. Insulators to Improve the Safety of Retroviral

Vectors for HIV Gene Therapy. *Biomedicines* 4, 4.

doi:10.3390/biomedicines4010004

23. Candotti, F., Shaw, K.L., Muul, L., Carbonaro, D., Sokolic, R., Choi, C., Schurman, S.H., Garabedian, E., Kesserwan, C., Jagadeesh, G.J., Fu, P.Y., Gschweng, E., Cooper, A., Tisdale, J.F., Weinberg, K.I., Crooks, G.M., Kapoor, N., Shah, A., Abdel-Azim, H., Yu, X.J., Smogorzewska, M., Wayne, A.S., Rosenblatt, H.M., Davis, C.M., Hanson, C., Rishi, R.G., Wang, X., Gjertson, D., Yang, O.O., Balamurugan, A., Bauer, G., Ireland, J.A., Engel, B.C., Podsakoff, G.M., Hershfield, M.S., Blaese, R.M., Parkman, R., Kohn, D.B., 2012. Gene therapy for adenosine deaminase-deficient severe combined immune deficiency: clinical comparison of retroviral vectors and treatment plans. *Blood* 120, 3635–3646. doi:10.1182/blood-2012-02-400937
24. Cant Barrett, K., Mendes, R.D., Smits, W.K., van Helsing-van Wijk, Y.M., Pieters, R., Meijerink, J.P.P., 2016. Lentiviral gene transfer into human and murine hematopoietic stem cells: size matters. *BMC Res Notes* 9, 6. doi:10.1186/s13104-016-2118-z
25. Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P., Luc, S., Kurita, R., Nakamura, Y., Fujiwara, Y., Maeda, T., Yuan, G.-C., Zhang, F., Orkin, S.H., Bauer, D.E., 2015. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* 527, 192–197. doi:10.1038/nature15521
26. Carbonaro, D.A., Zhang, L., Jin, X., Montiel-Equihua, C., Geiger, S., Carmo, M., Cooper, A., Fairbanks, L., Kaufman, M.L., Sebire, N.J., Hollis, R.P., Blundell,

- M.P., Senadheera, S., Fu, P.-Y., Sahaghian, A., Chan, R.Y., Wang, X., Cornetta, K., Thrasher, A.J., Kohn, D.B., Gaspar, H.B., 2014. Preclinical demonstration of lentiviral vector-mediated correction of immunological and metabolic abnormalities in models of adenosine deaminase deficiency. *Molecular Therapy* 22, 607–622. doi:10.1038/mt.2013.265
27. Carroll, D., 2016. Genome editing: progress and challenges for medical applications. *Genome Medicine* 2016 8:1 8, 120. doi:10.1186/s13073-016-0378-9
28. Cartier, N., Aubourg, P., 2010. Hematopoietic Stem Cell Transplantation and Hematopoietic Stem Cell Gene Therapy in X-Linked Adrenoleukodystrophy. *Brain Pathology* 20, 857–862. doi:10.1111/j.1750-3639.2010.00394.x
29. Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C.C., Bougnères, P., Schmidt, M., Kalle, von, C., Fischer, A., Cavazzana-Calvo, M., Aubourg, P., 2012. Lentiviral Hematopoietic Cell Gene Therapy for X-Linked Adrenoleukodystrophy. *Methods in Enzymology* 507, 187–198. doi:10.1016/B978-0-12-386509-0.00010-7
30. Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C.C., Veres, G., Schmidt, M., Kutschera, I., Vidaud, M., Abel, U., Dal Cortivo, L., Caccavelli, L., Mahlaoui, N., Kiermer, V., Mittelstaedt, D., Bellesme, C., Lahlou, N., Lefrere, F., Blanche, S., Audit, M., Payen, E., Leboulch, P., l'Homme, B., Bougnères, P., Kalle, von, C., Fischer, A., Cavazzana-Calvo, M., Aubourg, P., 2009. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* 326, 818–823. doi:10.1126/science.1171242
31. Case, S.S., Price, M.A., Jordan, C.T., Yu, X.J., Wang, L., Bauer, G., Haas, D.L.,

- Xu, D., Stripecke, R., Naldini, L., Kohn, D.B., Crooks, G.M., 1999. Stable transduction of quiescent CD34+CD38- human hematopoietic cells by HIV-1-based lentiviral vectors. *Proceedings of the National Academy of Sciences* 96, 2988–2993. doi:10.1073/pnas.96.6.2988
32. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., Cavallesco, R., Gillet-Legend, B., Caccavelli, L., Sgarra, R., Maouche-Chrétien, L., Bernaudin, F., Girot, R., Dorazio, R., Mulder, G.-J., Polack, A., Bank, A., Soulier, J., Larghero, J., Kabbara, N., Dalle, B., Gourmel, B., Socie, G., Chrétien, S., Cartier, N., Aubourg, P., Fischer, A., Cornetta, K., Galacteros, F., Beuzard, Y., Gluckman, E., Bushman, F., Hacein-Bey-Abina, S., Leboulch, P., 2010. Transfusion independence and HMGA2 activation after gene therapy of human  $\beta$ -thalassaemia 467, 318–322. doi:10.1038/nature09328
33. Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., Voytas, D.F., 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* 39, 7879–7879. doi:10.1093/nar/gkr739
34. Cesana, D., Ranzani, M., Volpin, M., Bartholomae, C., Duros, C., Artus, A., Merella, S., Benedicenti, F., Sergi Sergi, L., Sanvito, F., Brombin, C., Nonis, A., Serio, C.D., Doglioni, C., Kalle, von, C., Schmidt, M., Cohen-Haguenaer, O., Naldini, L., Montini, E., 2014. Uncovering and Dissecting the Genotoxicity of Self-inactivating Lentiviral Vectors In Vivo. *Molecular Therapy* 22, 774–785. doi:10.1038/mt.2014.3

35. Chang, K.-H., Smith, S.E., Sullivan, T., Chen, K., Zhou, Q., West, J.A., Liu, M., Liu, Y., Vieira, B.F., Sun, C., Hong, V.P., Zhang, M., Yang, X., Reik, A., Urnov, F.D., Rebar, E.J., Holmes, M.C., Danos, O., Jiang, H., Tan, S., 2017. Long-Term Engraftment and Fetal Globin Induction upon BCL11A Gene Editing in Bone-Marrow-Derived CD34 + Hematopoietic Stem and Progenitor Cells. *Molecular Therapy - Methods & Clinical Development* 4, 137–148.  
doi:10.1016/j.omtm.2016.12.009
36. Chapman, J.R., Taylor, M.R.G., Boulton, S.J., 2012. Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Molecular Cell* 47, 497–510.  
doi:10.1016/j.molcel.2012.07.029
37. Chaudhury, S., Ayas, M., Rosen, C., Ma, M., Viqaruddin, M., Parikh, S., Kharbanda, S., Chiang, K.Y., Haight, A., Bhatia, M., Guilcher, G., Thompson, A., Shenoy, S., 2017. A Multicenter Retrospective Analysis Stressing the Importance of Long-Term Follow-Up after Hematopoietic Cell Transplantation for  $\beta$ -Thalassemia. *Biology of Blood and Marrow Transplantation* 23, 1695–1700.  
doi:10.1016/j.bbmt.2017.06.004
38. Cheshier, S.H., Morrison, S.J., Liao, X., Weissman, I.L., 1999. In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. *Proceedings of the National Academy of Sciences* 96, 3120–3125.  
doi:10.1073/pnas.96.6.3120
39. Chhabra, A., Ring, A.M., Weiskopf, K., Schnorr, P.J., Gordon, S., Le, A.C., Kwon, H.S., Ring, N.G., Volkmer, J., Ho, P.Y., Tseng, S., Weissman, I.L., Shizuru, J.A., 2016. Hematopoietic stem cell transplantation in immunocompetent hosts without

- radiation or chemotherapy. *Science Translational Medicine* 8, 351ra105–351ra105. doi:10.1126/scitranslmed.aae0501
40. Chiruvella, K.K., Liang, Z., Wilson, T.E., 2013. Repair of Double-Strand Breaks by End Joining. *Cold Spring Harbor Perspectives in Biology* 5, a012757–a012757. doi:10.1101/cshperspect.a012757
41. Clough, C., Wang, Y., Khan, I.F., Singh, S., Hung, K., Rawlings, D.J., 2016. 132. Targeting the BTK Locus in Primary Human Hematopoietic Cells with TALENs and AAV Donor Template. *Molecular Therapy* 24, S54. doi:10.1016/S1525-0016(16)32941-0
42. Cooke, K.R., Luznik, L., Sarantopoulos, S., Hakim, F.T., Jagasia, M., Fowler, D.H., van den Brink, M.R.M., Hansen, J.A., Parkman, R., Miklos, D.B., Martin, P.J., Paczesny, S., Vogelsang, G., Pavletic, S., Ritz, J., Schultz, K.R., Blazar, B.R., 2017. The Biology of Chronic Graft-versus-Host Disease: A Task Force Report from the National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease. *Biology of Blood and Marrow Transplantation* 23, 211–234. doi:10.1016/j.bbmt.2016.09.023
43. Cradick, T.J., Fine, E.J., Antico, C.J., Bao, G., 2013. CRISPR/Cas9 systems targeting  $\beta$ -globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res* 41, 9584–9592. doi:10.1093/nar/gkt714
44. Czechowicz, A., Kraft, D., Weissman, I.L., Bhattacharya, D., 2007. Efficient Transplantation via Antibody-Based Clearance of Hematopoietic Stem Cell Niches. *Science* 318, 1296–1299. doi:10.1126/science.1149726
45. Daneshbod-Skibba, G., Martin, J., Shahidi, N.T., 1980. Myeloid and Erythroid

- Colony Growth in Non-anaemic Patients with Fanconi's Anaemia. *British Journal of Haematology* 44, 33–38. doi:10.1111/j.1365-2141.1980.tb01181.x
46. De Ravin, S.S., Li, L., Wu, X., Choi, U., Allen, C., Koontz, S., Lee, J., Theobald-Whiting, N., Chu, J., Garofalo, M., Sweeney, C., Kardava, L., Moir, S., Viley, A., Natarajan, P., Su, L., Kuhns, D., Zarembek, K.A., Peshwa, M.V., Malech, H.L., 2017. CRISPR-Cas9 gene repair of hematopoietic stem cells from patients with X-linked chronic granulomatous disease. *Science Translational Medicine* 9, eaah3480. doi:10.1126/scitranslmed.aah3480
47. De Ravin, S.S., Reik, A., Liu, P.-Q., Li, L., Wu, X., Su, L., Raley, C., Theobald, N., Choi, U., Song, A.H., Chan, A., Pearl, J.R., Paschon, D.E., Lee, J., Newcombe, H., Koontz, S., Sweeney, C., Shivak, D.A., Zarembek, K.A., Peshwa, M.V., Gregory, P.D., Urnov, F.D., Malech, H.L., 2016a. Targeted gene addition in human CD34+ hematopoietic cells for correction of X-linked chronic granulomatous disease. *Nature Biotechnology* 34, 424–429. doi:10.1038/nbt.3513
48. De Ravin, S.S., Wu, X., Moir, S., Kardava, L., Anaya-O'Brien, S., Kwatema, N., Littel, P., Theobald, N., Choi, U., Su, L., Marquesen, M., Hilligoss, D., Lee, J., Buckner, C.M., Zarembek, K.A., O'Connor, G., McVicar, D., Kuhns, D., Throm, R.E., Zhou, S., Notarangelo, L.D., Hanson, I.C., Cowan, M.J., Kang, E., Hadigan, C., Meagher, M., Gray, J.T., Sorrentino, B.P., Malech, H.L., 2016b. Lentiviral hematopoietic stem cell gene therapy for X-linked severe combined immunodeficiency. *Science Translational Medicine* 8, 335ra57–335ra57. doi:10.1126/scitranslmed.aad8856

49. De Rijck, J., de Kogel, C., Demeulemeester, J., Vets, S., Ashkar, El, S., Malani, N., Bushman, F.D., Landuyt, B., Husson, S.J., Busschots, K., Gijsbers, R., Debyser, Z., 2013. The BET Family of Proteins Targets Moloney Murine Leukemia Virus Integration near Transcription Start Sites. *Cell Reports* 5, 886–894. doi:10.1016/j.celrep.2013.09.040
50. DeKelver, R.C., Choi, V.M., Moehle, E.A., Paschon, D.E., Hockemeyer, D., Meijsing, S.H., Sancak, Y., Cui, X., Steine, E.J., Miller, J.C., Tam, P., Bartsevich, V.V., Meng, X., Rupniewski, I., Gopalan, S.M., Sun, H.C., Pitz, K.J., Rock, J.M., Zhang, L., Davis, G.D., Rebar, E.J., Cheeseman, I.M., Yamamoto, K.R., Sabatini, D.M., Jaenisch, R., Gregory, P.D., Urnov, F.D., 2010. Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Research* 20, 1133–1142. doi:10.1101/gr.106773.110
51. Derse, D., Crise, B., Li, Y., Princler, G., Lum, N., Stewart, C., McGrath, C.F., Hughes, S.H., Munroe, D.J., Wu, X., 2007. Human T-Cell Leukemia Virus Type 1 Integration Target Sites in the Human Genome: Comparison with Those of Other Retroviruses. *Journal of Virology* 81, 6731–6741. doi:10.1128/JVI.02752-06
52. Dever, D.P., Bak, R.O., Reinisch, A., Camarena, J., Washington, G., Nicolas, C.E., Pavel-Dinu, M., Saxena, N., Wilkens, A.B., Mantri, S., Uchida, N., Hendel, A., Narla, A., Majeti, R., Weinberg, K.I., Porteus, M.H., 2016. CRISPR/Cas9  $\beta$ -globin gene targeting in human haematopoietic stem cells. *Nature* 539, 384–389. doi:10.1038/nature20134
53. DeWitt, M.A., Magis, W., Bray, N.L., Wang, T., Berman, J.R., Urbinati, F., Heo,

- S.-J., Mitros, T., Muñoz, D.P., Boffelli, D., Kohn, D.B., Walters, M.C., Carroll, D., Martin, D.I.K., Corn, J.E., 2016. Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Science Translational Medicine* 8, 360ra134–360ra134. doi:10.1126/scitranslmed.aaf9336
54. DiGiusto, D.L., Cannon, P.M., Holmes, M.C., Li, L., Rao, A., Wang, J., Lee, G., Gregory, P.D., Kim, K.A., Hayward, S.B., Meyer, K., Exline, C., Lopez, E., Henley, J., Gonzalez, N., Bedell, V., Stan, R., Zaia, J.A., 2016. Preclinical development and qualification of ZFN-mediated CCR5 disruption in human hematopoietic stem/progenitor cells. *Molecular Therapy - Methods & Clinical Development* 3, 16067. doi:10.1038/mtm.2016.67
55. Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., Naldini, L., 1998. A third-generation lentivirus vector with a conditional packaging system. *Journal of Virology* 72, 8463–8471.
56. Eichler, F., Duncan, C., Musolino, P.L., Orchard, P.J., De Oliveira, S., Thrasher, A.J., Armant, M., Dansereau, C., Lund, T.C., Miller, W.P., Raymond, G.V., Sankar, R., Shah, A.J., Sevin, C., Gaspar, H.B., Gissen, P., Amartino, H., Bratkovic, D., Smith, N.J.C., Paker, A.M., Shamir, E., O'Meara, T., Davidson, D., Aubourg, P., Williams, D.A., 2017. Hematopoietic Stem-Cell Gene Therapy for Cerebral Adrenoleukodystrophy. *N Engl J Med* NEJMoa1700554. doi:10.1056/NEJMoa1700554
57. Emery, D.W., 2011. The Use of Chromatin Insulators to Improve the Expression and Safety of Integrating Gene Transfer Vectors. *Human Gene Therapy* 22, 761–774. doi:10.1089/hum.2010.233

58. Emery, D.W., Liu, M., O, M., Dorschner, M.O., 2009. Gammaretroviral Vector Integration Occurs Overwhelmingly at Borders of DNase Hypersensitive Sites. *Molecular Therapy* 17, S144. doi:10.1016/S1525-0016(16)38726-3
59. Enssle, J., Trobridge, G.D., Keyser, K.A., Ironside, C., Beard, B.C., Kiem, H.-P., 2010. Stable Marking and Transgene Expression Without Progression to Monoclonality in Canine Long-Term Hematopoietic Repopulating Cells Transduced with Lentiviral Vectors. *Human Gene Therapy* 21, 397–403. doi:10.1089/hum.2009.076
60. Fares, I., Chagraoui, J., Gareau, Y., Gingras, S., Ruel, R., Mayotte, N., Csaszar, E., Knapp, D.J.H.F., Miller, P., Ngom, M., Imren, S., Roy, D.C., Watts, K.L., Kiem, H.P., Herrington, R., Iscove, N.N., Humphries, R.K., Eaves, C.J., Cohen, S., Marinier, A., Zandstra, P.W., Sauvageau, G., 2014. Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. *Science* 345, 1509–1512. doi:10.1126/science.1256337
61. Fuchs, E.J., 2015. HLA-haploidentical blood or marrow transplantation with high-dose, post-transplantation cyclophosphamide. *Bone Marrow Transplant* 50, S31–S36. doi:10.1038/bmt.2015.92
62. Gaspar, H.B., Buckland, K., Rivat, C., Himoudi, N., Gilmour, K., Booth, C., Cornetta, K., Kohn, D.B., Paruzynski, A., Schmidt, M., Thrasher, A.J., 2014. Immunological and Metabolic Correction After Lentiviral Vector Mediated Haematopoietic Stem Cell Gene Therapy for ADA Deficiency. *Molecular Therapy* 22, S106. doi:10.1016/S1525-0016(16)35289-3
63. Gaspar, H.B., Cooray, S., Gilmour, K.C., Parsley, K.L., Zhang, F., Adams, S.,

Bjorkegren, E., Bayford, J., Brown, L., Davies, E.G., Veys, P., Fairbanks, L., Bordon, V., Petropoulou, T., Kinnon, C., Thrasher, A.J., 2011. Hematopoietic Stem Cell Gene Therapy for Adenosine Deaminase-Deficient Severe Combined Immunodeficiency Leads to Long-Term Immunological Recovery and Metabolic Correction. *Science Translational Medicine* 3, 97ra80–97ra80.

doi:10.1126/scitranslmed.3002716

64. Gelsinger, P., Shamo, A.E., 2008. Eight Years after Jesse's Death, Are Human Research Subjects Any Safer? *Hastings Center Report* 38, 25–27.

doi:10.1353/hcr.2008.0022

65. Gennari, F., Lopes, L., Verhoeyen, E., Marasco, W., Collins, M.K., 2009. Single-Chain Antibodies That Target Lentiviral Vectors to MHC Class II on Antigen-Presenting Cells. *Human Gene Therapy* 20, 554–562.

doi:10.1089/hum.2008.189

66. Genovese, P., Schirotti, G., Escobar, G., Di Tomaso, T., Firrito, C., Calabria, A., Moi, D., Mazzieri, R., Bonini, C., Holmes, M.C., Gregory, P.D., van der Burg, M., Gentner, B., Montini, E., Lombardo, A., Naldini, L., 2014. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* 510, 235–240.

doi:10.1038/nature13420

67. Girard-Gagnepain, A., Amirache, F., Costa, C., Levy, C., Frecha, C., Fusil, F., Negre, D., Lavillette, D., Cosset, F.L., Verhoeyen, E., 2014. Baboon envelope pseudotyped LVs outperform VSV-G-LVs for gene transfer into early-cytokine-stimulated and resting HSCs. *Blood* 124, 1221–1231. doi:10.1182/blood-2014-

02-558163

68. Giri, N., Batista, D.L., Alter, B.P., Stratakis, C.A., 2007. Endocrine Abnormalities in Patients with Fanconi Anemia. *The Journal of Clinical Endocrinology & Metabolism* 92, 2624–2631. doi:10.1210/jc.2007-0135
69. Goessling, W., Allen, R.S., Guan, X., Jin, P., Uchida, N., Dovey, M., Harris, J.M., Metzger, M.E., Bonifacino, A.C., Stroncek, D., Stegner, J., Armant, M., Schlaeger, T., Tisdale, J.F., Zon, L.I., Donahue, R.E., North, T.E., 2011. Prostaglandin E2 Enhances Human Cord Blood Stem Cell Xenotransplants and Shows Long-Term Safety in Preclinical Nonhuman Primate Transplant Models. *Cell Stem Cell* 8, 445–458. doi:10.1016/j.stem.2011.02.003
70. Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G.F., Chin, L., 2016. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Reports* 14, 1555–1566. doi:10.1016/j.celrep.2016.01.019
71. Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., Asnafi, V., MacIntyre, E., Dal Cortivo, L., Radford, I., Brousse, N., Sigaux, F., Moshous, D., Hauer, J., Borkhardt, A., Belohradsky, B.H., Wintergerst, U., Velez, M.C., Leiva, L., Sorensen, R., Wulffraat, N., Blanche, S., Bushman, F.D., Fischer, A., Cavazzana-Calvo, M., 2008. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* 118, 3132–3142. doi:10.1172/JCI35700
72. Hacein-Bey-Abina, S., Kalle, von, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntyre, E., Radford, I., Villeval, J.-L., Fraser, C.C., Cavazzana-Calvo, M., Fischer, A., 2003. A Serious Adverse Event after Successful Gene Therapy for

X-Linked Severe Combined Immunodeficiency. *N Engl J Med* 348, 255–256.  
doi:10.1056/NEJM200301163480314

73. Hacein-Bey-Abina, S., Le Deist, F., Carlier, F., Bouneaud, C., Hue, C., De Villartay, J.-P., Thrasher, A.J., Wulffraat, N., Sorensen, R., Dupuis-Girod, S., Fischer, A., Davies, E.G., Kuis, W., Leiva, L., Cavazzana-Calvo, M., 2002. Sustained Correction of X-Linked Severe Combined Immunodeficiency by ex Vivo Gene Therapy. *N Engl J Med* 346, 1185–1193. doi:10.1056/NEJMoa012616
74. Hacein-Bey-Abina, S., Pai, S.-Y., Gaspar, H.B., Armant, M., Berry, C.C., Blanche, S., Bleesing, J., Blondeau, J., de Boer, H., Buckland, K.F., Caccavelli, L., Cros, G., De Oliveira, S., Fernandez, K.S., Guo, D., Harris, C.E., Hopkins, G., Lehmann, L.E., Lim, A., London, W.B., van der Loo, J.C.M., Malani, N., Male, F., Malik, P., Marinovic, M.A.L., McNicol, A.-M., Moshous, D., Neven, B., Oleastro, M.A., Picard, C., Ritz, J., Rivat, C., Schambach, A., Shaw, K.L., Sherman, E.A., Silberstein, L.E., Six, E., Touzot, F., Tsytsykova, A., Xu-Bayford, J., Baum, C., Bushman, F.D., Fischer, A., Kohn, D.B., Filipovich, A.H., Notarangelo, L.D., Cavazzana, M., Williams, D.A., Thrasher, A.J., 2014. A Modified  $\gamma$ -Retrovirus Vector for X-Linked Severe Combined Immunodeficiency. *N Engl J Med* 371, 1407–1417. doi:10.1056/NEJMoa1404588
75. Hanenberg, H., Xiao, X.L., Dilloo, D., Hashino, K., Kato, I., Williams, D.A., 1996. Colocalization of retrovirus and target cells on specific fibronectin fragments increases genetic transduction of mammalian cells. *Nat Med* 2, 876–882.  
doi:10.1038/nm0896-876
76. Hendel, A., Bak, R.O., Clark, J.T., Kennedy, A.B., Ryan, D.E., Roy, S., Steinfeld,

- I., Lunstad, B.D., Kaiser, R.J., Wilkens, A.B., Bacchetta, R., Tsalenko, A., Dellinger, D., Bruhn, L., Porteus, M.H., 2015. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nature Biotechnology* 33, 985–989. doi:10.1038/nbt.3290
77. Heyer, W.-D., Ehmsen, K.T., Liu, J., 2010. Regulation of Homologous Recombination in Eukaryotes. *Annu. Rev. Genet.* 44, 113–139. doi:10.1146/annurev-genet-051710-150955
78. Hoban, M.D., Cost, G.J., Mendel, M.C., Romero, Z., Kaufman, M.L., Joglekar, A.V., Ho, M., Lumaquin, D., Gray, D., Lill, G.R., Cooper, A.R., Urbinati, F., Senadheera, S., Zhu, A., Liu, P.Q., Paschon, D.E., Zhang, L., Rebar, E.J., Wilber, A., Wang, X., Gregory, P.D., Holmes, M.C., Reik, A., Hollis, R.P., Kohn, D.B., 2015. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood* 125, 2597–2604. doi:10.1182/blood-2014-12-615948
79. Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKolver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B., Meng, X., Miller, J.C., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., Jaenisch, R., 2009. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nature Biotechnology* 27, 851–857. doi:10.1038/nbt.1562
80. Hofmann, A., Kessler, B., Ewerling, S., Kabermann, A., Brem, G., Wolf, E., Pfeifer, A., 2006. Epigenetic Regulation of Lentiviral Transgene Vectors in a Large Animal Model. *Molecular Therapy* 13, 59–66. doi:10.1016/j.ymthe.2005.07.685

81. Hoggatt, J., Singh, P., Sampath, J., Pelus, L.M., 2009. Prostaglandin E2 enhances hematopoietic stem cell homing, survival, and proliferation. *Blood* 113, 5444–5455. doi:10.1182/blood-2009-01-201335
82. Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G.M., Kohn, D.B., Gregory, P.D., Holmes, M.C., Cannon, P.M., 2010. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nature Biotechnology* 28, 839–847. doi:10.1038/nbt.1663
83. Hope, T., 2002. Improving the Post-Transcriptional Aspects of Lentiviral Vectors, in: *Lentiviral Vectors, Current Topics in Microbiology and Immunology*. Springer, Berlin, Heidelberg, Berlin, Heidelberg, pp. 179–189. doi:10.1007/978-3-642-56114-6\_9
84. Howe, S.J., Mansour, M.R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempfski, H., Brugman, M.H., Pike-Overzet, K., Chatters, S.J., de Ridder, D., Gilmour, K.C., Adams, S., Thornhill, S.I., Parsley, K.L., Staal, F.J.T., Gale, R.E., Linch, D.C., Bayford, J., Brown, L., Quaye, M., Kinnon, C., Ancliff, P., Webb, D.K., Schmidt, M., Kalle, von, C., Gaspar, H.B., Thrasher, A.J., 2008. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.* 118, 3143–3150. doi:10.1172/JCI35798
85. Hsieh, M.M., Wu, C.J., Tisdale, J.F., 2011. In mixed hematopoietic chimerism, the donor red cells win. *Haematologica* 96, 13–15. doi:10.3324/haematol.2010.035576

86. Hubbard, N., Hagin, D., Sommer, K., Song, Y., Khan, I., Clough, C., Ochs, H.D., Rawlings, D.J., Scharenberg, A.M., Torgerson, T.R., 2016. Targeted gene editing restores regulated CD40L function in X-linked hyper-IgM syndrome. *Blood* 127, 2513–2522. doi:10.1182/blood-2015-11-683235
87. Hustedt, N., Durocher, D., 2016. The control of DNA repair by the cell cycle. *Nat Cell Biol* 19, 1–9. doi:10.1038/ncb3452
88. Jasin, M., 1996. Genetic manipulation of genomes with rare-cutting endonucleases. *Trends in Genetics* 12, 224–228. doi:10.1016/0168-9525(96)10019-6
89. Kim, S., Kim, D., Cho, S.W., Kim, J., Kim, J.S., 2014. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research* 24, 1012–1019. doi:10.1101/gr.171322.113
90. Kim, Y.G., Cha, J., Chandrasegaran, S., 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences* 93, 1156–1160. doi:10.1073/pnas.93.3.1156
91. Kondo, M., Wagers, A.J., Manz, M.G., Prohaska, S.S., Scherer, D.C., Beilhack, G.F., Shizuru, J.A., Weissman, I.L., 2003. Biology of Hematopoietic Stem Cells and Progenitors: Implications for Clinical Application. *Annual Review of Immunology* 21, 759–806. doi:10.1146/annurev.immunol.21.120601.141007
92. Kumar, M., Keller, B., Makalou, N., Sutton, R.E., 2004. Systematic Determination of the Packaging Limit of Lentiviral Vectors. *Human Gene Therapy* 12, 1893–1905. doi:10.1089/104303401753153947
93. Lewis, P.F., Emerman, M., 1994. Passage through mitosis is required for

oncoretroviruses but not for the human immunodeficiency virus. *Viral Gene Techniques, Methods in Molecular Genetics* 68, 510–516. doi:10.1016/s1067-2389(06)80044-0

94. Li, H., Haurigot, V., Doyon, Y., Li, T., Wong, S.Y., Bhagwat, A.S., Malani, N., Anguela, X.M., Sharma, R., Ivanciu, L., Murphy, S.L., Finn, J.D., Khazi, F.R., Zhou, S., Paschon, D.E., Rebar, E.J., Bushman, F.D., Gregory, P.D., Holmes, M.C., High, K.A., 2011. In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* 475, 217–221. doi:10.1038/nature10177
95. Li, L., Krymskaya, L., Wang, J., Henley, J., Rao, A., Cao, L.-F., Tran, C.-A., Torres-Coronado, M., Gardner, A., Gonzalez, N., Kim, K., Liu, P.-Q., Hofer, U., Lopez, E., Gregory, P.D., Liu, Q., Holmes, M.C., Cannon, P.M., Zaia, J.A., DiGiusto, D.L., 2013. Genomic Editing of the HIV-1 Coreceptor CCR5 in Adult Hematopoietic Stem and Progenitor Cells Using Zinc Finger Nucleases. *Molecular Therapy* 21, 1259–1269. doi:10.1038/mt.2013.65
96. Lieber, M.R., Ma, Y., Pannicke, U., Schwarz, K., 2003. Mechanism and regulation of human non-homologous DNA end-joining. *Nature Reviews Molecular Cell Biology* 4, 712–720. doi:10.1038/nrm1202
97. Lin, S., Stahl, B.T., Alla, R.K., Doudna, J.A., 2014. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* 3, 1314. doi:10.7554/eLife.04766
98. Lis, R., Karrasch, C.C., Poulos, M.G., Kunar, B., Redmond, D., Duran, J.G.B., Badwe, C.R., Schachterle, W., Ginsberg, M., Xiang, J., Tabrizi, A.R., Shido, K., Rosenwaks, Z., Elemento, O., Speck, N.A., Butler, J.M., Scandura, J.M., Rafii,

- S., 2017. Conversion of adult endothelium to immunocompetent haematopoietic stem cells. *Nature* 545, 439–445. doi:10.1038/nature22326
99. Lombardo, A., Pietro Genovese, Beausejour, C.M., Colleoni, S., Lee, Y.-L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D., Holmes, M.C., Naldini, L., 2007. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nature Biotechnology* 25, 1298–1306. doi:10.1038/nbt1353
100. Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., Church, G.M., 2013a. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology* 31, 833–838. doi:10.1038/nbt.2675
101. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., 2013b. RNA-Guided Human Genome Engineering via Cas9. *Science* 339, 823–826. doi:10.1126/science.1232033
102. Mandal, P.K., Ferreira, L.M.R., Collins, R., Meissner, T.B., Boutwell, C.L., Friesen, M., Vrbanac, V., Garrison, B.S., Stortchevoi, A., Bryder, D., Musunuru, K., Brand, H., Tager, A.M., Allen, T.M., Talkowski, M.E., Rossi, D.J., Cowan, C.A., 2014. Efficient Ablation of Genes in Human Hematopoietic Stem and Effector Cells using CRISPR/Cas9. *Cell Stem Cell* 15, 643–652. doi:10.1016/j.stem.2014.10.004
103. Maruyama, T., Dougan, S.K., Truttmann, M.C., Bilate, A.M., Ingram, J.R., Ploegh, H.L., 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nature Biotechnology*

- 33, 538–542. doi:10.1038/nbt.3190
104. Masiuk, K.E., Brown, D., Laborada, J., Hollis, R.P., Urbinati, F., Kohn, D.B., 2017. Improving Gene Therapy Efficiency Through the Enrichment of Human Hematopoietic Stem Cells. *Molecular Therapy* 25, 2163–2175. doi:10.1016/j.ymthe.2017.05.023
105. Matreyek, K., Engelman, A., 2013. Viral and Cellular Requirements for the Nuclear Entry of Retroviral Preintegration Nucleoprotein Complexes. *Viruses* 2013, Vol. 5, Pages 2483-2511 5, 2483–2511. doi:10.3390/v5102483
106. Metzger, L., Iliakis, G., 2009. Kinetics of DNA Double-strand Break Repair Throughout the Cell Cycle as Assayed by Pulsed Field Gel Electrophoresis in CHO Cells. *International Journal of Radiation Biology* 59, 1325–1339. doi:10.1080/09553009114551201
107. Miller, D.G., Adam, M.A., Miller, A.D., 1990. Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection. *Mol. Cell. Biol.* 10, 4239–4242. doi:10.1128/MCB.10.8.4239
108. Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., Dulay, G.P., Hua, K.L., Ankoudinova, I., Cost, G.J., Urnov, F.D., Zhang, H.S., Holmes, M.C., Zhang, L., Gregory, P.D., Rebar, E.J., 2010. A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology* 29, 143–148. doi:10.1038/nbt.1755
109. Millington, M., Arndt, A., Boyd, M., Applegate, T., Shen, S., 2009. Towards a Clinically Relevant Lentiviral Transduction Protocol for Primary Human CD34(+) Hematopoietic Stem/Progenitor Cells. *PLoS ONE* 4, e6461.

doi:10.1371/journal.pone.0006461

110. Mitchell, R.S., Beitzel, B.F., Schroder, A.R.W., Shinn, P., Chen, H., Berry, C.C., Ecker, J.R., Bushman, F.D., 2004. Retroviral DNA Integration: ASLV, HIV, and MLV Show Distinct Target Site Preferences. *PLoS Biol* 2, e234.

doi:10.1371/journal.pbio.0020234

111. Miyoshi, H., Smith, K.A., Mosier, D.E., Verma, I.M., Torbett, B.E., 1999. Transduction of human CD34+ cells that mediate long-term engraftment of NOD/SCID mice by HIV vectors. *Science* 283, 682–686.

doi:10.1126/science.283.5402.682

112. Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M.H., Schambach, A., Charrier, S., Galy, A., Thrasher, A.J., Bueren, J., Baum, C., 2009. Insertional Transformation of Hematopoietic Cells by Self-inactivating Lentiviral and Gammaretroviral Vectors. *Molecular Therapy* 17, 1919–1928.

doi:10.1038/mt.2009.179

113. Moreno-Carranza, B., Gentsch, M., Stein, S., Schambach, A., Santilli, G., Rudolf, E., Ryser, M.F., Haria, S., Thrasher, A.J., Baum, C., Brenner, S., Grez, M., 2008. Transgene optimization significantly improves SIN vector titers, gp91phox expression and reconstitution of superoxide production in X-CGD cells. *Gene Ther* 16, 111–118. doi:10.1038/gt.2008.143

114. Morgan, R., Boyerinas, B., 2016. Genetic Modification of T Cells.

*Biomedicines* 4, 9. doi:10.3390/biomedicines4020009

115. Muccio, L., Bertaina, A., Falco, M., Pende, D., Meazza, R., Lopez-Botet, M., Moretta, L., Locatelli, F., Moretta, A., Chiesa, M.D., 2016. Analysis of memory-

- like natural killer cells in human cytomegalovirus-infected children undergoing +T and B cell-depleted hematopoietic stem cell transplantation for hematological malignancies. *Haematologica* 101, 371–381.  
doi:10.3324/haematol.2015.134155
116. Nakade, S., Tsubota, T., Sakane, Y., Kume, S., Sakamoto, N., Obara, M., Daimon, T., Sezutsu, H., Yamamoto, T., Sakuma, T., Suzuki, K.-I.T., 2014. Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat Comms* 5, ncomms6560. doi:10.1038/ncomms6560
117. Naldini, L., Blomer, U., Gally, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., Trono, D., 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272, 263–267.  
doi:10.1126/science.272.5259.263
118. Ng, Y.Y., Baert, M.R.M., Pike-Overzet, K., Rodijk, M., Brugman, M.H., Schambach, A., Baum, C., Hendriks, R.W., van Dongen, J.J.M., Staal, F.J.T., 2010. Correction of B-cell development in Btk-deficient mice using lentiviral vectors with codon-optimized human BTK. *Leukemia* 24, 1617–1630.  
doi:10.1038/leu.2010.140
119. Nienhuis, A.W., Dunbar, C.E., Sorrentino, B.P., 2006. Genotoxicity of Retroviral Integration In Hematopoietic Cells. *Molecular Therapy* 13, 1031–1049.  
doi:10.1016/j.ymthe.2006.03.001
120. North, T.E., Goessling, W., Walkley, C.R., Lengerke, C., Kopani, K.R., Lord, A.M., Weber, G.J., Bowman, T.V., Jang, I.-H., Grosser, T., FitzGerald, G.A.,

- Daley, G.Q., Orkin, S.H., Zon, L.I., 2007. Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. *Nature* 447, 1007–1011.  
doi:10.1038/nature05883
121. Notta, F., Doulatov, S., Laurenti, E., Poepl, A., Jurisica, I., Dick, J.E., 2011. Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science* 333, 218–221. doi:10.1126/science.1201219
122. Orkin, S.H., Reilly, P., 2016. Paying for future success in gene therapy. *Science* 352, 1059–1061. doi:10.1126/science.aaf4770
123. Orthwein, A., Noordermeer, S.M., Wilson, M.D., Landry, S., Enchev, R.I., Sherker, A., Munro, M., Pinder, J., Salsman, J., Dellaire, G., Xia, B., Peter, M., Durocher, D., 2015. A mechanism for the suppression of homologous recombination in G1 cells. *Nature* 528, 422–426. doi:10.1038/nature16142
124. Ott, M.G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., Glimm, H., Kühlcke, K., Schilz, A., Kunkel, H., Naundorf, S., Brinkmann, A., Deichmann, A., Fischer, M., Ball, C., Pilz, I., Dunbar, C., Du, Y., Jenkins, N.A., Copeland, N.G., Lüthi, U., Hassan, M., Thrasher, A.J., Hoelzer, D., Kalle, von, C., Seger, R., Grez, M., 2006. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat Med* 12, 401–409. doi:10.1038/nm1393
125. Palchaudhuri, R., Saez, B., Hoggatt, J., Schajnovitz, A., Sykes, D.B., Tate, T.A., Czechowicz, A., Kfoury, Y., Ruchika, F., Rossi, D.J., Verdine, G.L., Mansour, M.K., Scadden, D.T., 2016. Non-genotoxic conditioning for hematopoietic stem cell transplantation using a hematopoietic-cell-specific

internalizing immunotoxin. *Nature Biotechnology* 34, 738–745.

doi:10.1038/nbt.3584

126. Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K.M., Gregg, A., Noggle, S., Tessier-Lavigne, M., 2016. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533, 125–129. doi:10.1038/nature17664
127. Passegué, E., Wagers, A.J., Giuriato, S., Anderson, W.C., Weissman, I.L., 2005. Global analysis of proliferation and cell cycle gene expression in the regulation of hematopoietic stem and progenitor cell fates. *J Exp Med* 202, 1599–1611. doi:10.1084/jem.20050967
128. Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K.A., Liu, O., Wang, N., Lee, G., Bartsevich, V.V., Lee, Y.-L., Guschin, D.Y., Rupniewski, I., Waite, A.J., Carpenito, C., Carroll, R.G., S Orange, J., Urnov, F.D., Rebar, E.J., Ando, D., Gregory, P.D., Riley, J.L., Holmes, M.C., June, C.H., 2008. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nature Biotechnology* 26, 808–816. doi:10.1038/nbt1410
129. Peters, C., Charnas, L.R., Tan, Y., Ziegler, R.S., Shapiro, E.G., DeFor, T., Grewal, S.S., Orchard, P.J., Abel, S.L., Goldman, A.I., Ramsay, N.K.C., Dusenbery, K.E., Loes, D.J., Lockman, L.A., Kato, S., Aubourg, P.R., Moser, H.W., Krivit, W., 2004. Cerebral X-linked adrenoleukodystrophy: the international hematopoietic cell transplantation experience from 1982 to 1999. *Blood* 104, 881–888. doi:10.1182/blood-2003-10-3402
130. Petrillo, C., Cesana, D., Piras, F., Bartolaccini, S., Naldini, L., Montini, E.,

- Kajaste-Rudnitski, A., 2015. Cyclosporin A and Rapamycin Relieve Distinct Lentiviral Restriction Blocks in Hematopoietic Stem and Progenitor Cells. *Molecular Therapy* 23, 352–362. doi:10.1038/mt.2014.193
131. Pietras, E.M., Warr, M.R., Passegué, E., 2011. Cell cycle regulation in hematopoietic stem cells. *J Cell Biol* 195, 709–720. doi:10.1083/jcb.201102131
132. Pinder, J., Salsman, J., Dellaire, G., 2015. Nuclear domain “knock-in” screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Res* 43, 9379–9392. doi:10.1093/nar/gkv993
133. Ramezani, A., Hawley, T.S., Hawley, R.G., 2008. Combinatorial Incorporation of Enhancer Blocking Components of the Chicken  $\beta$ -Globin 5'HS4 and Human T-Cell Receptor  $\alpha/\delta$  BEAD-1 Insulators in Self-Inactivating Retroviral Vectors Reduces their Genotoxic Potential. *Stem Cells* 26, 3257–3266. doi:10.1634/stemcells.2008-0258
134. Richardson, C.D., Kazane, K.R., Feng, S.J., Bray, N.L., Schaefer, A.J., Floor, S., Corn, J., 2017. CRISPR-Cas9 Genome Editing In Human Cells Works Via The Fanconi Anemia Pathway. *bioRxiv*. doi:10.1101/136028
135. Sadelain, M., May, C., Rivella, S., Callegari, J., Heller, G., Gaensler, K.M.L., Luzzatto, L., 2000. Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. *Nature* 406, 82–86. doi:10.1038/35017565
136. Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.-I.T., Yamamoto, T., 2015. MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCH

- systems. *Nat Protoc* 11, 118–133. doi:10.1038/nprot.2015.140
137. Santoni de Sio, F.R., 2006. Proteasome activity restricts lentiviral gene transfer into hematopoietic stem cells and is down-regulated by cytokines that enhance transduction. *Blood* 107, 4257–4265. doi:10.1182/blood-2005-10-4047
138. Saydaminova, K., Ye, X., Wang, H., Richter, M., Ho, M., Chen, H., Xu, N., Kim, J.-S., Papapetrou, E., Holmes, M.C., Gregory, P.D., Palmer, D., Ng, P., Ehrhardt, A., Lieber, A., 2015. Efficient genome editing in hematopoietic stem cells with helper-dependent Ad5/35 vectors expressing site-specific endonucleases under microRNA regulation. *Molecular Therapy - Methods & Clinical Development* 2, 14057. doi:10.1038/mtm.2014.57
139. Schambach, A., Wodrich, H., Hildinger, M., Bohne, J., Krusslich, H.-G., Baum, C., 2000. Context Dependence of Different Modules for Posttranscriptional Enhancement of Gene Expression from Retroviral Vectors. *Molecular Therapy* 2, 435–445. doi:10.1006/mthe.2000.0191
140. Schirotti, G., Genovese, P., Castiello, M.C., Capo, V., Gregory, P.D., Holmes, M.C., Sitia, G., Villa, A., Lombardo, A., Naldini, L., 2015. 481. Targeted Genome Editing in Mouse Hematopoietic Stem/Progenitor Cells (HSPC) To Model Gene Correction of SCID-X1. *Molecular Therapy* 23, S191. doi:10.1016/S1525-0016(16)34090-4
141. Shaw, K.L., Garabedian, E., Mishra, S., Barman, P., Davila, A., Carbonaro, D., Shupien, S., Silvin, C., Geiger, S., Nowicki, B., Smogorzewska, E.M., Brown, B., Wang, X., De Oliveira, S., Choi, Y., Ikeda, A., Terrazas, D., Fu, P.-Y., Yu, A., Fernandez, B.C., Cooper, A.R., Engel, B., Podsakoff, G., Balamurugan, A.,

- Anderson, S., Muul, L., Jagadeesh, G.J., Kapoor, N., Tse, J., Moore, T.B., Purdy, K., Rishi, R., Mohan, K., Skoda-Smith, S., Buchbinder, D., Abraham, R.S., Scharenberg, A., Yang, O.O., Cornetta, K., Gjertson, D., Hershfield, M., Sokolic, R., Candotti, F., Kohn, D.B., 2017. Clinical efficacy of gene-modified stem cells in adenosine deaminase-deficient immunodeficiency. *J. Clin. Invest.* 127, 1689–1699. doi:10.1172/JCI90367
142. Shrivastav, M., De Haro, L.P., Nickoloff, J.A., 2008. Regulation of DNA double-strand break repair pathway choice. *Cell Res* 18, 134–147. doi:10.1038/cr.2007.111
143. Spitz, F., Furlong, E.E.M., 2012. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* 13, 613–626. doi:10.1038/nrg3207
144. Srivastava, M., Nambiar, M., Sharma, S., Karki, S.S., Goldsmith, G., Hegde, M., Kumar, S., Pandey, M., Singh, R.K., Ray, P., Natarajan, R., Kelkar, M., De, A., Choudhary, B., Raghavan, S.C., 2012. An Inhibitor of Nonhomologous End-Joining Abrogates Double-Strand Break Repair and Impedes Cancer Progression. *Cell* 151, 1474–1487. doi:10.1016/j.cell.2012.11.054
145. Stein, S., Ott, M.G., Schultze-Strasser, S., Jauch, A., Burwinkel, B., Kinner, A., Schmidt, M., Krämer, A., Schwäble, J., Glimm, H., Koehl, U., Preiss, C., Ball, C., Martin, H., Göhring, G., Schwarzwaelder, K., Hofmann, W.-K., Karakaya, K., Tchatchou, S., Yang, R., Reinecke, P., Kühlcke, K., Schlegelberger, B., Thrasher, A.J., Hoelzer, D., Seger, R., Kalle, von, C., Grez, M., 2010. Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat Med* 16, 198–204.

doi:10.1038/nm.2088

146. Sugimura, R., Jha, D.K., Han, A., Soria-Valles, C., da Rocha, E.L., Lu, Y.-F., Goettel, J.A., Serrao, E., Rowe, R.G., Malleshaiah, M., Wong, I., Sousa, P., Zhu, T.N., Ditadi, A., Keller, G., Engelman, A.N., Snapper, S.B., Doulatov, S., Daley, G.Q., 2017. Haematopoietic stem and progenitor cells from human pluripotent stem cells. *Nature* 545, 432–438. doi:10.1038/nature22370
147. Sutton, R.E., Reitsma, M.J., Uchida, N., Brown, P.O., 1999. Transduction of Human Progenitor Hematopoietic Stem Cells by Human Immunodeficiency Virus Type 1-Based Vectors Is Cell Cycle Dependent. *Viral Vectors for Gene Therapy* 73, 3649–3660. doi:10.1385/1-59259-304-6:467
148. Suzuki, K., Tsunekawa, Y., Hernandez-Benitez, R., Wu, J., Zhu, J., Kim, E.J., Hatanaka, F., Yamamoto, M., Araoka, T., Li, Z., Kurita, M., Hishida, T., Li, M., Aizawa, E., Guo, S., Chen, S., Goebel, A., Soligalla, R.D., Qu, J., Jiang, T., Fu, X., Jafari, M., Esteban, C.R., Berggren, W.T., Lajara, J., Nuñez-Delicado, E., Guillen, P., Campistol, J.M., Matsuzaki, F., Liu, G.-H., Magistretti, P., Zhang, K., Callaway, E.M., Zhang, K., Belmonte, J.C.I., 2016. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* 540, 144–149. doi:10.1038/nature20565
149. Tebas, P., Stein, D., Tang, W.W., Frank, I., Wang, S.Q., Lee, G., Spratt, S.K., Surosky, R.T., Giedlin, M.A., Nichol, G., Holmes, M.C., Gregory, P.D., Ando, D.G., Kalos, M., Collman, R.G., Binder-Scholl, G., Plesa, G., Hwang, W.-T., Levine, B.L., June, C.H., 2014. Gene Editing of CCR5 in Autologous CD4 T Cells of Persons Infected with HIV. *N Engl J Med* 370, 901–910.

doi:10.1056/NEJMoa1300662

150. Thom, R.E., Ouma, A.A., Zhou, S., Chandrasekaran, A., Lockey, T., Greene, M., De Ravin, S.S., Moayeri, M., Malech, H.L., Sorrentino, B.P., Gray, J.T., 2009. Efficient construction of producer cell lines for a SIN lentiviral vector for SCID-X1 gene therapy by concatemeric array transfection. *Blood* 113, 5104–5110.  
doi:10.1182/blood-2008-11-191049
151. Van Tendeloo, V.F.I., 2001. Highly efficient gene delivery by mRNA electroporation in human hematopoietic cells: superiority to lipofection and passive pulsing of mRNA and to electroporation of plasmid cDNA for tumor antigen loading of dendritic cells. *Blood* 98, 49–56. doi:10.1182/blood.V98.1.49
152. Van Trung Chu, Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., Kühn, R., 2015. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature Biotechnology* 33, 543–548. doi:10.1038/nbt.3198
153. Verhoeven, E., Wiznerowicz, M., Olivier, D., Izac, B., Trono, D., Dubart-Kupperschmitt, A., Cosset, F.-L., 2005. Novel lentiviral vectors displaying “early-acting cytokines” selectively promote survival and transduction of NOD/SCID repopulating human hematopoietic stem cells. *Blood* 106, 3386–3395.  
doi:10.1182/blood-2004-12-4736
154. Vink, C.A., Counsell, J.R., Perocheau, D.P., Karda, R., Buckley, S.M.K., Brugman, M.H., Galla, M., Schambach, A., McKay, T.R., Waddington, S.N., Howe, S.J., 2017. Eliminating HIV-1 Packaging Sequences from Lentiviral Vector Proviruses Enhances Safety and Expedites Gene Transfer for Gene Therapy.

Molecular Therapy 25, 1790–1804. doi:10.1016/j.ymthe.2017.04.028

155. Walters, M.C., 2015. Update of hematopoietic cell transplantation for sickle cell disease. *Current Opinion in Hematology* 22, 227–233.  
doi:10.1097/MOH.000000000000136
156. Wang, C.X., Sather, B.D., Wang, X., Adair, J., Khan, I., Singh, S., Lang, S., Adams, A., Curinga, G., Kiem, H.P., Miao, C.H., Rawlings, D.J., Torbett, B.E., 2014. Rapamycin relieves lentiviral vector transduction resistance in human and mouse hematopoietic stem cells. *Blood* 124, 913–923. doi:10.1182/blood-2013-12-546218
157. Wang, J., Exline, C.M., DeClercq, J.J., Llewellyn, G.N., Hayward, S.B., Li, P.W.-L., Shivak, D.A., Surosky, R.T., Gregory, P.D., Holmes, M.C., Cannon, P.M., 2015. Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nature Biotechnology* 33, 1256–1263. doi:10.1038/nbt.3408
158. West, A.G., Fraser, P., 2005. Remote control of gene transcription. *Hum Mol Genet* 14, R101–R111. doi:10.1093/hmg/ddi104
159. Wilson, J.M., 2009. Lessons learned from the gene therapy trial for ornithine transcarbamylase deficiency. *Molecular Genetics and Metabolism* 96, 151–157.  
doi:10.1016/j.ymgme.2008.12.016
160. Woychik, R.P., Lyons, R.H., Post, L., Rottman, F.M., 1984. Requirement for the 3' flanking region of the bovine growth hormone gene for accurate polyadenylation. *Proceedings of the National Academy of Sciences* 81, 3944–3948. doi:10.1073/pnas.81.13.3944

161. Wright, A.V., Nuñez, J.K., Doudna, J.A., 2016. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell* 164, 29–44. doi:10.1016/j.cell.2015.12.035
162. Wu, H.-W., Hsu, R.-C., Lin, C.-C., Hwang, S.-M., Lee, G.-B., 2010. An integrated microfluidic system for isolation, counting, and sorting of hematopoietic stem cells. *Biomicrofluidics* 4, 024112. doi:10.1063/1.3454767
163. Wu, X., 2003. Transcription Start Regions in the Human Genome Are Favored Targets for MLV Integration. *Science* 300, 1749–1751. doi:10.1126/science.1083413
164. Xue, X., Pech, N.K., Shelley, W.C., Srour, E.F., Yoder, M.C., Dinauer, M.C., 2010. Antibody targeting KIT as pretransplantation conditioning in immunocompetent mice. *Blood* 116, 5419–5422. doi:10.1182/blood-2010-07-295949
165. Yang, D., Scavuzzo, M.A., Chmielowiec, J., Sharp, R., Bajic, A., Borowiak, M., 2016. Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Sci Rep* 6, 861. doi:10.1038/srep21264
166. Yao, S., Sukonnik, T., Kean, T., Bharadwaj, R.R., Pasceri, P., Ellis, J., 2004. Retrovirus Silencing, Variegation, Extinction, and Memory Are Controlled by a Dynamic Interplay of Multiple Epigenetic Modifications. *Molecular Therapy* 10, 27–36. doi:10.1016/j.ymthe.2004.04.007
167. Yu, S.F., Ruden, von, T., Kantoff, P.W., Garber, C., Seiberg, M., Ruther, U., Anderson, W.F., Wagner, E.F., Gilboa, E., 1986. Self-inactivating retroviral

vectors designed for transfer of whole genes into mammalian cells. Proceedings of the National Academy of Sciences 83, 3194–3198.

doi:10.1073/pnas.83.10.3194

168. Zhou, S., Fatima, S., Ma, Z., Wang, Y.-D., Lu, T., Janke, L.J., Du, Y., Sorrentino, B.P., 2016. Evaluating the Safety of Retroviral Vectors Based on Insertional Oncogene Activation and Blocked Differentiation in Cultured Thymocytes. *Molecular Therapy* 24, 1090–1099. doi:10.1038/mt.2016.55
169. Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., Trono, D., 1998. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *Journal of Virology* 72, 9873–9880.
170. Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L., Trono, D., 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nature Biotechnology* 15, 871–875. doi:10.1038/nbt0997-871

## **Web resources**

1. bluebird bio, Inc., (2017). Published online September 06, 2017.  
<http://www.businesswire.com/news/home/20170906005645/en/>
2. European Medicines Agency, (2016). Published online April 01, 2016.  
[http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Press\\_release/2016/04/WC500204145.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2016/04/WC500204145.pdf).

**Chapter 2:  $\beta$ -Globin Locus Control Region Core Sequences Driving Expression of Anti-Sickling Globin Ameliorates Disease Phenotype in a Mouse Model of Sickle Cell**

## **Abstract**

$\beta$ -globin lentiviral vectors ( $\beta$ -LV) have faced challenges in clinical translation for gene therapy of sickle cell disease (SCD) due to low titer and sub-optimal gene transfer to hematopoietic stem and progenitor cells (HSPC). To overcome the challenge of preserving efficacious expression while increasing vector performance, we used published genomic and epigenomic data available through ENCODE to redefine enhancer element boundaries of the  $\beta$ -globin Locus Control Region (LCR) to construct novel ENCODE core sequences. These novel LCR elements were used to design a  $\beta$ -LV of reduced proviral length, termed CoreGA-AS3-FB, produced at higher titers and possessing superior gene transfer to HSPCs when compared to the full-length parental  $\beta$ -LV at equal multiplicity of infection. At low vector copy number, vectors containing the ENCODE core sequences were capable of reversing the sickle phenotype in a mouse model of SCD. These studies provide a  $\beta$ -LV that will be beneficial for gene therapy of SCD by significantly reducing the cost of vector production and extending the vector supply.

## **Introduction**

Sickle Cell Disease (SCD) is caused by a single nucleotide mutation in the  $\beta$ -globin gene, which results in the formation of abnormally charged hemoglobin (HbS) protein in red blood cells (RBC)(Ingram, 1957). Vaso-occlusion is a major clinical feature of SCD and is caused by sickling of RBCs due to aggregation of HbS under low oxygen conditions(Francis and Johnson, 1991). While two Food and Drug Administration approved drug treatments for SCD exist, both Hydroxyurea and L-

Glutamine only modestly improve overall health of afflicted patients even when adherence is strictly maintained(Charache et al., 1995; Niihara et al., 2018).

As an alternative therapy allogeneic hematopoietic stem cell transplantation (HSCT) has the potential to cure patients afflicted with SCD, however identification of suitable HLA-matched donors remains a significant obstacle(Justus et al., 2015). Autologous HSCT combined with gene therapy is a more contemporary approach that entails the genetic modification of a patient's own CD34+ hematopoietic stem and progenitor cells (HSPC). By replacing a patient's native HSPC population with genetically modified HSPCs harboring a normal or modified  $\beta$ -globin gene (such as Lenti/ $\beta$ AS3-FB(Urbinati et al., 2017)), SCD can be permanently cured without risks of graft vs host disease or graft rejection that accompanies the more traditional strategy of allogeneic HSCT(Ribeil et al., 2017). Preliminary data from clinical trials for hemoglobin disorders have indeed demonstrated a benefit from autologous HSCT with gene therapy(Kanter et al., 2016; Rasko et al., 2017; Ribeil et al., 2017; Thompson et al., 2018; Walters et al., 2017).

Low vector titer has represented a major barrier toward the effective advancement of therapeutic  $\beta$ -globin lentiviral vectors ( $\beta$ -LV) in clinical applications. Typically,  $\beta$ -LVs are produced at 10- to 100-fold lower titers when compared to LV with shorter genomes carrying small regulatory elements and cDNAs. In addition to the problem of low titer, clinical grade  $\beta$ -LV preparations (typically concentrated 100x-500x(Cooper et al., 2011)) showed reduced transduction efficiency for primary human HSPC and did not provide dose-dependent transduction of HSPCs with increasing vector concentrations (Kohn 2018 – unpublished data). The combinations of lower titer

and subpar transduction efficiency mean that large volumes of vector need to be produced per patient dose, significantly adding to the clinical expense.

Poor gene transfer to HSPCs has also been a challenge in gene therapy trials and has limited the therapeutic benefit of gene therapy for patients with  $\beta$ -Thalassemia and SCD(Cavazzana-Calvo et al., 2010; Rasko et al., 2017). Although the issues of gene transfer to HSPCs has been partially mitigated through the advent of transduction enhancers(Anastasov et al., 2016; Beschorner et al., 2018; Delville et al., 2018; Petrillo et al., 2018) the purported use of transduction enhancers to achieve adequate VCNs in clinic(Kanter et al., 2016; Ribeil et al., 2017; Walters et al., 2017) does not lower the cost of vector per patient dose when clinical vectors with inherently low infectivity are used.

One predominant factor that diminishes  $\beta$ -LV titer is the relatively large size of these vector genomes (e.g. 6-9 kb) due to the use of the human  $\beta$ -globin genomic sequences (promoter, exons, introns, 3' untranslated region) and the  $\beta$ -globin Locus Control Region (LCR) DNase 1 hypersensitive site (HS) elements that augment expression. Another critical factor limiting vector performance is the inherent complexity of the human genomic elements; the  $\beta$ -globin genomic sequences and LCR are transcribed in reverse orientation relative to the vector genomic transcript (to retain intronic enhancers) and may contain problematic sequences that lead to non-functional vector genomes (e.g. cryptic polyadenylation signals).

The Globe-AS3-FB  $\beta$ -LV produced by Ferrari and co-workers(Miccio et al., 2008) is ~2.7 Kb smaller than the Lenti/ $\beta$ AS3-FB  $\beta$ -LV we have used(Romero et al., 2013) (as well as other commonly used  $\beta$ -LVs(May et al., 2002; Negre et al., 2015; Pawliuk et al.,

2001; Perumbeti et al., 2009)) and lacks the LCR HS4 sequences, the WPRE (Woodchuck Hepatitis Virus Post Transcriptional Regulatory Element), and possesses a smaller 3'UTR segment of the  $\beta$ -globin gene and a larger deletion of  $\beta$ -globin intron 2. It is unknown how sequence differences contribute to Globe-AS3-FB's superior packaging and transduction efficiency when compared to Lenti/ $\beta$ AS3-FB(Urbinati, 2018) or how these individual sequence differences influence  $\beta$ -globin expression.

Thus, we set out to determine if the detrimental differences in titer and gene transfer of Lenti/ $\beta$ AS3-FB when compared to Globe-AS3-FB are due to proviral length or specific sequence elements. To reduce proviral size we employed current genomic and epigenomic databases to identify and redefine functional elements present within the human genomic sequences of the LCR in Lenti/ $\beta$ AS3-FB to produce a composite LCR. A series of modified vectors were made with the aims of increasing titer and transduction efficiency while retaining high-level lineage-specific expression. Modified constructs were compared head-to-head and results were used to guide construction of a reduced length vector with superior titer and gene transfer, and with sufficient expression of the anti-sickling  $\beta^{\text{AS3}}$ -globin gene to induce hematologic correction of the sickle phenotype in the "Townes" mouse model of SCD.

## **RESULTS**

### **Proviral Length Influences Packaging and Transduction Efficiency**

Previous reports have demonstrated that titer and transduction efficiency of lentiviral vectors inversely correlate with proviral length(Canté-Barrett et al., 2016; Kumar et al., 2001). To determine whether the differences of performances between

Lenti/ $\beta$ AS3-FB and Globe-AS3-FB were due to vector length (as opposed to the presence of specific adverse sequences), a series of derivatives that reflect each difference between the two constructs were cloned and tested (Figure 1A). When constructs were packaged and titered head-to-head, an inverse correlation between titer and vector genome length was observed (Figure 1B). When human CD34+ HSPCs were transduced at  $2 \times 10^7$  transduction units (TU)/mL (multiplicity of infection [MOI] 20), cultured for 2 weeks under myeloid culture conditions<sup>ii</sup> and vector copy number (VCN) plotted as a function of vector genome length, an inverse relationship was also observed between genome length and VCN (Figure 1C) with donor-to-donor variation in transduction efficiency diminishing the strength of the correlation. Notably, the GLOBE-based constructs showed reduced transduction efficiency for their size, relative to the other vector.

Direct evidence that vector genome length was the major factor impacting vector performance was obtained through simply deleting the HS4 region from Lenti/ $\beta$ AS3-FB, as it was a major difference between Lenti/ $\beta$ AS3-FB and the higher titer vector Globe-AS3-FB. A two-fold increase in both titer ( $p < 0.01$ ) and gene transduction efficiency to CD34+ HSPCs ( $p < 0.05$ ) was observed when the 1.1Kb HS4 element was removed from Lenti/ $\beta$ AS3-FB. These gains in performance were not sequence-related, as replacing HS4 with a similarly-sized 1.1Kb stuffer sequence from the human *HPRT* gene intron 1 similarly reduced titer and infectivity (Figure 1D and 1E). As expected, the removal of

---

<sup>ii</sup> Transduced HSPCs were cultured under myeloid culture conditions (as opposed to erythroid culture conditions) as we have shown that short-term myeloid culture more stringently reflects the VCNs seen in bone marrow after xenotransplantation immune-deficient mouse hosts (Romero et al., 2013).

HS4 predictably decreased expression of  $\beta^{\text{AS3}}$ -globin(Lisowski and Sadelain, 2007; Navas et al., 2001) (Figure 1F)

Additional sequence deletions (> 1.1 Kb) offered increasing gains in titer and infectivity suggesting that length influenced vector performance more strongly than the presence of discreet sequences (additional analyses examining derivatives of Lenti/ $\beta$ AS3-FB are provided in the supplementary results). Taken together, these data suggested that the titer and infectivity of Lenti/ $\beta$ AS3-FB could be improved through shortening vector length. We hypothesized that a decrease in vector length of 2-3Kb would be required to generate a construct produced at three- to four-fold higher raw titer that would have superior gene transfer to HSPCs when compared to Lenti/ $\beta$ AS3-FB at equivalent MOI.

## Redefining the Putative Boundaries of HS Elements

Vector length reduction was achieved by redefining the boundaries of the LCR HS (2,3 and 4) elements of Lenti/ $\beta$ AS3-FB by using published genomic and epigenomic data available through ENCODE (Accessible via the UCSC Genome Browser(The ENCODE Project Consortium, 2012)). The “Open Chromatin” track sets, which combine, 1). histone modification, 2), histone displacement, 3). transcription factor binding, 4). DNase1-accessibility, and 5). sequence conservation were used to generate new boundaries for the HS sequences of Lenti/ $\beta$ AS3-FB. These novel defined HS sequences, termed “ENCODE core sequences” (EC2, ~420bp; EC3, ~340bp; EC4, ~410bp), were used to replace the HS sequences present within Lenti/ $\beta$ AS3-FB (~3.6 Kb reduced to ~1.2 Kb) to generate a reduced length construct termed “Core-AS3-FB”. We then added a 360bp fusion element to facilitate position-independent expression of  $\beta^{AS3}$ -globin by combining a 202bp murine Gata1 gene HS element with a 143bp Human Ankyrin gene element(Romero et al., 2015) to create CoreGA-AS3-FB (Figure 2A).

The  $\beta$ -LV constructs were packaged in parallel using HEK 293T cells and titered head-to-head on HT-29 cells(Cooper et al., 2011). From five independent packaging experiments, titers of the Globe-AS3-FB, Core-AS3-FB, and CoreGA-AS3-FB  $\beta$ -LVs were on average about three-fold higher than that of Lenti/ $\beta$ AS3-FB ( $p < 0.0001$ ), presumably due to their reduced proviral lengths in comparison to Lenti/ $\beta$ AS3-FB (Figure 2B).

To evaluate the efficiency of transduction of Core-AS3-FB and CoreGA-AS3-FB in comparison to Lenti/ $\beta$ AS3-FB and GLOBE-AS3-FB, human CD34+ HSPCs isolated

from bone marrow (BM) of healthy donors were transduced at  $1 \times 10^7$  TU/mL (MOI 10) and cultured for 14 days under myeloid differentiation conditions. From three independent experiments, the gene transfer efficiency of Core-AS3-FB was on average about three-fold higher when compared to Lenti/ $\beta$ AS3-FB and Globe-AS3-FB ( $p < 0.01$ ), and the further addition of the fusion element to produce CoreGA-AS3-FB did not negate gains in gene transfer efficiency achieved by Core-AS3 ( $p < 0.01$ ), (Figure 2C).

To confirm that CoreGA-AS3-FB retained superior transduction efficiency when compared to Lenti/ $\beta$ AS3-FB and GLOBE-AS3-FB at equal MOI ( $1 \times 10^6$  TU/ml), infectivity was assessed at the clonal level. Transduced CD34+ HSPCs were plated in a methylcellulose-based medium to allow growth of CFUs. After 14 days of culture, colonies were scored manually by morphologic criteria to determine lineage differentiation (BFU-E, CFU-E, CFU-GM, and CFU-GEMM). The VCNs of bulk transduced cells used to seed the CFU assay were 0.23, 0.29, and 0.67 for Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, and CoreGA-AS3-FB, respectively, reflecting the higher infectivity of the CoreGA-AS3-FB vector. Colonies were grouped by their VCN value as determined by ddPCR. Colonies with VCN  $> 0.3$  were considered PCR-positive for the presence of vector and further classified by VCN (0.3-1, 1-2, 2-4, or  $> 4$ ). The percentages of PCR-positive colonies seen for CoreGA-AS3-FB were significantly higher than for Lenti/ $\beta$ AS3-FB and Globe-AS3-FB ( $p < 0.004$ ; chi-squared test) reaching 51%, 18%, and 19%, respectively (Figure 2D). Among the positive colonies, CoreGA-AS3-FB achieved greater numbers of integrations per colony when compared to Lenti/ $\beta$ AS3-FB (0.3-1 VCN;  $p < 0.01$ , 1-2 VCN,  $p < 0.05$ ; 2-4 VCN,  $p < 0.005$ ; chi-squared test). These results demonstrate that when assessed at comparable MOI, CoreGA-AS3-FB was able

to transduce greater numbers of CD34+ HSPCs with more integrant per cells when compared to current clinical vectors.

To determine if the newly designed “Core” constructs retained the ability to achieve therapeutic levels of  $\beta^{\text{AS3}}$ -globin expression, *in vitro* erythroid differentiation studies were performed using BM CD34+ HSPCs from healthy donors. After HSPCs were transduced at equal MOI,  $\beta^{\text{AS3}}$ -globin and total  $\beta$ -globin RNA transcript levels were measured by reverse transcription (RT) droplet digital polymerase chain reaction (ddPCR) 14 days after transduction and *in vitro* differentiation. The percentages of  $\beta^{\text{AS3}}$ -globin RNA transcripts over total  $\beta$ -globin RNA transcript were normalized to VCN for nine independent experiments. Expression levels normalized by VCN of Core-AS3-FB and CoreGA-AS3-FB were 0.5-fold ( $p < 0.0001$ ), and 0.4-fold ( $p < 0.001$ ) lower than that of Lenti/ $\beta$ AS3-FB, respectively (Figure 2E).

To examine the linear relationship between VCN and % $\beta^{\text{AS3}}$ -globin transcript levels, HSPCs derived from multiple healthy donors were transduced at different MOIs and subjected to *in vitro* erythroid differentiation for 14 days, followed by measurement of expression of  $\beta^{\text{AS3}}$ -globin and VCN. Globe-AS3-FB had slightly higher levels of expression than did CoreGA-AS3-FB, (as shown by the increased slope of the fitted line), but this difference failed to reach significance. As Core-AS3-FB and CoreGA-AS3-FB were able to transduce HSPCs to higher VCN than the other vectors, total expression levels also increased in a linear fashion to compensate for the decreased expression per vector genome (Figure 2F).

Thus, *in vivo* studies comparing CoreGA-AS3-FB to Lenti/ $\beta$ AS3-FB and Globe-AS3-FB were conducted in the “Townes” mouse model of SCD(Levasseur et al., 2003;

2004; Ryan et al., 1997) to evaluate long term expression of  $\beta^{\text{AS3}}$ -globin and to determine if CoreGA-AS3-FB could ameliorate the sequelae related to SCD.

### ***In vivo* Analysis of Peripheral Blood from SCD Mouse Model**

The “Townes” mouse model of SCD was used to evaluate the hematologic correction potential of CoreGA-AS3-FB. While we hypothesized that the superior transduction efficiency of CoreGA-AS3-FB would compensate for its decreased expression per vector genome when compared to Lenti/ $\beta^{\text{AS3}}$ -FB at equal MOI, we chose to perform a direct comparison of CoreGA-AS3-FB to Lenti/ $\beta^{\text{AS3}}$ -FB and Globe-AS3-FB (Figure 3A) at equivalent VCN. Using constructs under conditions where transduction efficiencies were comparable (achieved by using vectors at different MOI) enabled a more appropriate evaluation for hematologic correction at low copy number. An outline of the experiment is provided in Figure 3B.

Lineage-depleted bone marrow cells were obtained from homozygous  $\beta^{\text{S}}/\beta^{\text{S}}$  donor mice and pre-stimulated for 24-hours. Cells were transduced with different amounts of vector to achieve equivalent copy numbers and delivered by retro-orbital injection into lethally irradiated recipients (Pep Boy/J [CD45.1 congenic]) 24-hours after transduction. Table 1 provides the concentrations of vector used to transduce murine lineage depleted (Lin-) BM cells. Two independent experiments were conducted and *in vitro* VCN was determined from the transduced cell product 14 days after transduction (shown in Table 1).

Engraftment by flow cytometry for CD45.2 (+) donor cells, transduction efficiency (VCN by ddPCR),  $\beta^{\text{AS3}}$ -globin expression (by RT-ddPCR), blood hemoglobin

concentration and RBC counts and hemoglobin composition by high-performance liquid chromatography (HPLC) were measured on peripheral blood (PB) samples acquired 4, 16 and 20 weeks post-transplantation. At week 20, mice were euthanized, and BM cells collected to measure engraftment, transduction efficiency, and  $\beta^{\text{AS3}}$ -globin expression. Mice with BM donor engraftment <97% at week 20 were excluded from analyses to prevent inclusion of artefactually-corrected mice (as  $\geq 4\%$  residual WT recipient RBCs would mask the adverse pathophysiology induced by  $\beta^{\text{S}}/\beta^{\text{S}}$  donor cells (Levasseur et al., 2003). At the time of euthanasia, BM cells taken from two mice from each arm were used to seed a methylcellulose-based colony forming unit (CFU) assay; after two weeks growth, individual primitive erythroid progenitor cell colonies (BFU-E) were isolated from the CFU dishes and assayed to measure VCN and  $\beta^{\text{AS3}}$ -globin expression. Peripheral blood cells were also collected at time of euthanasia and used to quantify the percentages of sickled erythrocytes present within the peripheral circulation.

As intended, the average gene transfer efficiency seen in PB cells did not significantly differ among experimental arms and the average PB VCNs were similar to those of *in vitro* cultured cells before transplantation. Moreover, gene marking of cells remained stable across all three time-points indicating that cells with long-term engraftment potential were stably transduced and able to contribute to hematopoiesis (Figure 4A).

Quantification of Hb $\beta^{\text{AS3}}$  tetramers in peripheral blood lysates was accomplished using HPLC. Supplementary figure 4A provides representative HPLC chromatograms of globin tetramers in PB obtained at time of euthanasia from representative mice. At 20 weeks post transplantation, the average levels of Hb $\beta^{\text{AS3}}$ /total hemoglobin tetramers

were 37.4%, 28.8%, and 19.1% for Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, and CoreGA-AS3-FB, respectively (Figure 4B).

To compare the difference in normalized expression between each experimental arm, %Hb $\beta^{AS3}$  was normalized to PB VCN for each mouse and plotted. An average of 18.3, 13.4, and 8.8 %Hb $\beta^{AS3}$ / VCN was seen for Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, and CoreGA-AS3-FB, respectively (Figure 4C). Supplementary table 1 provides additional characterizations of globin tetramers in PB lysates generated by HPLC for weeks 4, 16, and 20.

The Hb levels seen at time of euthanasia were significantly higher when comparing recipients of BM transduced with Lenti/ $\beta$ AS3-FB ( $p < 0.0001$ ), Globe-AS3-FB ( $p < 0.05$ ), and CoreGA-AS3-FB ( $p < 0.05$ ) to recipients of mock-transduced BM cells. The Hb levels of mice that received mock-transduced cells were 8.5 g/dL on average while the Hb levels of mice that received BM transduced with Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, or CoreGA-AS3-FB were 10.7, 10.2, and 9.9 g/dL on average, respectively (Figure 4D).

RBC counts were also significantly higher for recipients of Lenti/ $\beta$ AS3-FB ( $p < 0.001$ ), Globe-AS3-FB ( $p < 0.01$ ), and CoreGA-AS3-FB ( $p < 0.001$ ) transduced BM cells compared to recipients of mock-transduced BM cells. The RBC counts of mice that received mock-transduced cells were  $6.0 \times 10^6$  cells/ $\mu$ L on average, while the RBC counts of mice that received Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, or CoreGA-AS3-FB were 8.4, 8.5, and  $8.3 \times 10^6$  cells/ $\mu$ L on average, respectively (Figure 4E). Notably, mice that received mock transduced BM cells developed progressive anemia as noted by decreasing Hb levels and RBC counts observed in PB over time.

Lastly, the quantity of sickled RBCs present within PB at time of euthanasia from a subset of representative mice was measured using the ImageStream X Mark II Imaging Flow Cytometer (ISX). Equal volumes of peripheral blood and 2% sodium metabisulfite solution were mixed together and incubated at 37°C for 30 minutes and then analyzed by ISX. Peripheral blood taken from “Townes” homozygous  $\beta^S/\beta^S$  SCD mice or mice transplanted with mock-transduced HSPCs displayed averages of 41.0% and 34.3% RBC sickling, respectively. Mice transplanted with Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, or CoreGA-AS3-FB transduced HSPCs displayed a significant decrease in the percentages of sickled cells present within peripheral blood ( $p < 0.0001$ ). The average amount of sickling seen for Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, and CoreGA-AS3-FB were 7.5%, 5.76%, and 11.5%, respectively (Figure 4F).

### ***In vivo* Analysis of Bone Marrow $\beta^{AS3}$ -Globin Expression from SCD Mouse Model**

Bone marrow cells were collected at time of euthanasia and used to measure engraftment, transduction efficiency, and  $\beta^{AS3}$ -globin expression. As expected, average BM VCN was similar among all experimental arms and matched the average PB VCN seen at euthanasia (Figure 5A).

To determine if CoreGA-AS3-FB retained the ability to achieve therapeutic levels of expression in BM cells,  $\beta^{AS3}$ -globin and total  $\beta$ -globin RNA transcript levels were measured by RT ddPCR. Average levels of % $\beta^{AS3}$ -globin were 21.5%, 15.1%, and 11.4% for Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, and CoreGA-AS3-FB, respectively (Figure 5B). To compare the difference in expression per vector genome between all treatment arms, % $\beta^{AS3}$ -globin expression was normalized to BM VCN for each mouse and plotted

in Figure 5C. An average of 9.1, 7.2, and 6.1 % $\beta^{\text{AS3}}$ -globin/total  $\beta$ -globin/VCN was seen for Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, and CoreGA-AS3-FB, respectively.

Persistence of long-term  $\beta^{\text{AS3}}$ -globin expression was determined using a methylcellulose-based CFU assay. Bone marrow was taken from two mice from each treatment arm at time of euthanasia and used to seed methylcellulose supplemented with recombinant cytokines that promoted primitive erythroid progenitor cell growth. Two independent experiments were performed. The VCNs of the bulk BM cells used to seed methylcellulose at the time of euthanasia are provided in Table 2. Colonies recognized as BFU-E were plucked 14 days after seeding and both DNA and RNA obtained from the same colony. The VCN of each colony was determined by ddPCR and colonies then grouped by their VCN value. The percentages of PCR-positive colonies seen for Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, and CoreGA-AS3-FB were 86%, 87%, and 84%, respectively (Figure 5D).

Persistence of long-term  $\beta^{\text{AS3}}$ -globin expression was determined by measuring % $\beta^{\text{AS3}}$ -globin for each BFU-E. The VCNs of PCR-positive BFU-Es with detectable endogenous  $\beta$ -globin expression were plotted in Figure 5E. While all experimental arms displayed populations of BFU-Es with >20% $\beta^{\text{AS3}}$ -globin, average levels of % $\beta^{\text{AS3}}$ -globin/total  $\beta$ -globin mRNA were 22.9%, 24.2%, and 15.8% for Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, and CoreGA-AS3-FB, respectively (Figure 5F). When % $\beta^{\text{AS3}}$ -globin expression values of individual CFUs were normalized to their corresponding VCNs, averages of 11.2%, 12.2%, and 7.1%  $\beta^{\text{AS3}}$ -globin/VCN were observed for Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, and CoreGA-AS3-FB, respectively (Figure 5G). Taken together, these data demonstrate

that long-term and persistent expression was observed in each experimental arm and differences in expression observed in bulk BM cells were reflected at the clonal level.

### **Incorporating ENCODE Core Sequences into a Shorter $\beta$ -LV Expression Cassette**

The expression studies of CoreGA-AS3-FB were performed to ensure that exchanging the original LCR sequence for the ENCODE core sequences could still provide sufficient expression to ameliorate disease sequelae in the Towne's mouse model of SCD. Since reduction in vector size was determined to be the largest factor influencing titer and gene transfer, we made further modifications to the CoreGA-AS3-FB to remove approximately 2 kb of sequence from the provirus. The main deletions were of the WPRE sequence,  $\beta$ -globin 3'UTR and IVS2 (matching those found in the Globe-AS3-FB). The resulting ~4.7 kb UV-AS3-FB "Ultimate Vector" (shown in Figure 6A) has a titer that is 3-fold higher than CoreGA-AS3-FB (Figure 6B). Gene transfer is also improved 3-fold compared to CoreGA-AS3-FB (Figure 6C) and the expression remains identical (Figure 6D).

### **Discussion**

The requirements to be met for a gene therapy vector to be considered as a candidate for clinical translation include high-titer, efficient and stable proviral transmission, high-level, lineage-specific, and persistent expression, and safe integration profile. The results presented herein demonstrate that some of the major hurdles that have historically restricted clinical translation of  $\beta$ -globin gene therapy vectors for treatment of SCD can be surmounted through vector engineering.

One predominant factor that diminishes  $\beta$ -LV titer is suspected to be the relatively large size of these vectors due to the use of the human  $\beta$ -globin genomic sequences and LCR elements. Both Kumar *et al.* and Barrett *et al.*, demonstrated that titers of LVs decrease proportionally to the length of insert irrespective of sequence (Canté-Barrett *et al.*, 2016; Kumar *et al.*, 2001). We confirmed that vector length was the major factor contributing to the difference in titer observed between Lenti/ $\beta$ AS3-FB and Globe-AS3-FB. Moreover, introduction of additional sequence deletions to Lenti/ $\beta$ AS3-FB offered increasing gains in titer and infectivity confirming that vector length strongly influenced titer and infectivity. Although the dissection revealed that vector length could be reduced by  $\sim$ 1kb by removing redundant sequences (see supplementary information), that level of size reduction fell short of our 2-3kb target.

The critical minimal segments of the  $\sim$ 20kb human  $\beta$ -globin LCR needed for high-level, position-independent, erythroid expression of  $\beta$ -globin were elucidated using transgenic mouse models that, at the time, relied on a limited repertoire of restriction enzymes to introduce sequence deletions to the hypersensitive sites (Fedosyuk and Peterson, 2007; Grosveld *et al.*, 1987; Peterson *et al.*, 1996; 2012; Ryan *et al.*, 1989; Tanimoto *et al.*, 2003). The findings of these studies informed the development of current first generation  $\beta$ -globin lentiviral vectors, which were engineered to contain specific LCR fragments that confer erythroid-specific expression of  $\beta$ -globin (Collis *et al.*, 1990; Forrester *et al.*, 1989; Fraser *et al.*, 1990; 1993; May *et al.*, 2000). Given the methods that were deployed to define the boundaries of the LCR HS elements, we hypothesized that these fragments included core sequences that bestow erythroid specific expression, but also some extraneous sequences that could be eliminated with

minimal decrease in enhancer activity. Therefore, emphasis was given to develop a strategy for reducing the length of the LCR HS elements recognizing potential risks for reducing overall expression per vector genome.

When designing the new core HS elements, several genetic and epigenetic factors were considered including the sites binding histones associated with active genes, DNase1 hypersensitivity in erythroid lineages and sequence conservation. Using this design approach, the overall lengths of combined HS elements 2, 3, and 4 were reduced from 3.6kb to 1.2kb. Regulatory elements from the erythroid-specific murine *Gata1* and human *ANK1* genes were added to aid in position-independent expression.

When compared to the standard Lenti/ $\beta$ AS3-FB design, we found that the CoreGA-AS3-FB was produced at nearly three-fold higher titer ( $p < 0.0001$ ), possessed nearly three-fold higher gene transfer to CD34+ HSPCs at the same MOI as for standard vectors ( $p < 0.01$ ), and expressed nearly 60% of the amount of  $\beta^{AS3}$ -globin transcripts per vector genome in CD34+ HSPCs cultured under erythroid conditions ( $p < 0.001$ ). These observations of improved titer and transduction efficiency taken together spurred the continued study of CoreGA-AS3-FB to determine if the observed difference in expression when compared to Lenti/ $\beta$ AS3-FB was relevant to a clinical scenario.

Previous reports examining SCD patients who co-inherit the condition of hereditary persistence of fetal hemoglobin (where the distribution of HbF is thought to be pan-cellular) indicate that anti-sickling globin protein levels greater than 10-15% can ameliorate the clinical sequelae of SCD (Akinsheye et al., 2011; Forget, 1998; “Heritability of fetal hemoglobin, white cell count, and other clinical traits from a sickle

cell disease family cohort.," 2019; Pembrey et al., 1978; Platt et al., 1994). When CoreGA-AS3-FB was evaluated in a mouse model of SCD at an average BM VCN = 2,  $\beta^{\text{AS3}}$ -globin RNA comprised an average of 12% of the total  $\beta$ -globin transcripts and Hb $\beta^{\text{AS3}}$  tetramers made-up 18.5% of the total hemoglobin tetramers. The finding that the percentage of Hb  $\beta^{\text{AS3}}$  tetramers exceeded the relative abundance of  $\beta^{\text{AS3}}$ -globin mRNA is consistent with the increased affinity for  $\alpha$ -globin of the  $\beta^{\text{AS3}}$ -globin protein due to the G16D amino acid substitution ("ABNORMAL HUMAN HEMOGLOBINS. IX. CHEMISTRY OF HEMOGLOBIN J-BALTIMORE.," 1963; Levasseur et al., 2003). Thus, the *in vivo* percentages of  $\beta^{\text{AS3}}$ -globin RNA transcripts and protein seen for CoreGA-AS3-FB are at levels expected to be therapeutic, as the  $\beta^{\text{AS3}}$ -globin gene has been shown to have equivalent anti-sickling activity as  $\gamma$ -globin(Levasseur et al., 2004).

A methylcellulose-based CFU assay using BM cells obtained from mice 20-weeks after transplantation also revealed the presence of BFU-Es with >20% $\beta^{\text{AS3}}$ -globin in all experimental arms. This finding suggested that populations of corrected HSPCs capable of contributing to long-term erythropoiesis were established by each vector. More importantly, when murine HSPCs were transduced with CoreGA-AS3-FB at lower MOI (in comparison to those MOIs used to transduce cells with Lenti/ $\beta^{\text{AS3}}$ -FB or Globe-AS3-FB), CoreGA-AS3-FB transduced similar percentages of colony-forming progenitors resulting in similar bulk VCNs. These data suggest that CoreGA-AS3-FB has superior infectivity in comparison to Lenti/ $\beta^{\text{AS3}}$ -FB and Globe-AS3-FB, exemplified by its transduction efficiency even when constrained by using less vector.

The most critical pre-clinical test of any novel  $\beta$ -LV design being considered as a candidate for clinical translation is amelioration of hematologic parameters defining the

pathological phenotype of a SCD mouse model. While the amounts of  $\beta^{\text{AS3}}$ -globin expressed per VCN by CoreGA-AS3-FB were lower when compared to the other vectors, the levels were sufficient to result in physiologic improvements in the signs of SCD in the mice, with significant increases in both RBC count and hemoglobin content in comparison to recipients of mock-transduced BM.

When the quantity of sickled RBCs were measured in PB at time of euthanasia from a subset of representative mice, a higher amount of sickling was seen in PB of CoreGA-AS3-FB treated mice. This observation may have resulted from transplantation of a larger number of non-transduced cells into recipient mice; a result of using less vector.

Incorporation of the ENCODE core sequences from CoreGA-AS3-FB into the approximately 2 kb shorter UV-AS3-FB resulted in a superior titer and gene transfer. This new lentiviral vector design with reduced sizes of the LCR components and expression cassette should have significant advantages for clinical-scale production by providing far higher level of gene transfer to HSPC from lower amounts of vector. It may extend the benefits of autologous gene therapy for SCD by significantly reducing the cost of vector production and extending the vector supply.

## **Materials and Methods:**

### **Cloning and vector production**

The Lenti/BAS3-FB vector has been described previously (Romero et al., 2013). To introduce desired sequence deletions, sets of reverse-oriented primers were used to PCR amplify the Lenti/BAS3-FB plasmid backbone and the resultant linearized plasmids

were then phosphorylated and ligated. To seamlessly introduce sequences in place of deletions, gBlock gene fragments (Integrated DNA Technologies, Skokie IL) containing homology to the plasmid backbone were ordered and joined to linearized plasmids using the NEBuilder HiFi DNA Assembly kit (New England Biolabs, Ipswich, MA). All plasmids were sequence verified by Sanger sequencing (Laragen, Culver City, CA). The murine Gata1 element was from bp 438-639 (GenBank Sequence ID: U89137.1) and the human *ANK1* element is from bp 103987- 104129 (NCBI Reference Sequence: NG\_012820.2). The UV-AS3-FB vector was created by using NEBuilder HiFi DNA Assembly kit to fuse together PCR amplified EC2 from CoreGA-AS3-FB to gBlock gene fragments for the: cPPT,  $\beta$ -globin expression cassette, EC3, EC4, 3'LTR with polyadenylation enhancements (BGH polyA) and SV40 origin of replication.

Transient transfection of 293T cells using the third-generation packaging system(Dull et al., 1998) provided packaged virus particles. Viral supernatants were then directly used for titer determination or concentrated by tangential flow filtration, as described by Cooper et al(Cooper et al., 2011). Briefly, the HT-29 human colorectal carcinoma cell line was transduced with different dilutions of both raw and concentrated vectors. To calculate titers, cells were harvested and VCNs were determined by ddPCR approximately 60 hours post transduction.

### **BM CD34+ cell culture and transduction**

All bone marrow aspirates were obtained from voluntary healthy donors supplied by AllCells (Alameda, CA). BM mononuclear cells were isolated by Ficoll-Hypaque density gradient centrifugation. CD34+ HSPCs were enriched using CD34+ MicroBead

Kit (Miltenyi Biotec, Bergisch Gladbach, Germany). Enriched CD34<sup>+</sup> HSPCs were cryopreserved in Fetal Bovine Serum supplemented with 10% dimethyl sulfoxide (Sigma–Aldrich, St Louis, MO) in liquid nitrogen. Cells were thawed and plated on non-tissue culture-treated six-well plates pre-coated with RetroNectin (20 µg/ml, Takara Shuzo Co., Otsu Japan) at 1x10<sup>6</sup> cells/mL. Cells were pre-stimulated for 16-24 hours in X-Vivo 15 medium (Lonza, Basel, Switzerland) supplemented with 1x glutamine, penicillin, and streptomycin (Gemini Bio-Products, Sacramento, CA), human stem cell factor (50ng/mL), human Flt-3 ligand (50ng/mL), human thrombopoietin (50ng/mL), and human interleukin-3 (20ng/mL; all cytokines were acquired from PeproTech, Rocky Hill, NJ). Concentrated viral supernatants were used at various MOI to transduce CD34<sup>+</sup> HSPCs for 24 hours. These cells were washed, re-plated and cultured under myeloid or erythroid culture conditions; as described by Romero et al (Romero et al., 2013). On day 14 of culture, genomic DNA and/or mRNA was extracted from transduced cells.

### **DDPCR for VCN and %β<sup>AS3</sup> mRNA quantification**

Genomic DNA was extracted using PureLink™ Genomic DNA Mini Kit (Invitrogen, Waltham, MA). VCN was calculated by using probes SCD4 (Human Syndecan 4) as a reference and HIV-1 PSI as a target. ddPCR was carried out as described in Urbinati et al (Urbinati et al., 2017). RNeasy Plus Mini Kit (Qiagen, Valencia, CA) was used for RNA extraction followed by reverse transcription as described by Urbinati et al (Urbinati et al., 2017). Probes HBB<sup>TOTAL</sup> as a reference and Hbβ<sup>AS3</sup> as a target were used to generate droplets for digital droplet PCR (ddPCR), as described by Hindson et al (Hindson et al., 2011). Droplets were analyzed for absolute

quantification of the  $\beta^{\text{AS3}}$  gene expression normalized to the total B-globin gene expression.

### ***In vivo* experiment in SCD mouse model**

Bone marrow from eight to twelve-week-old homozygous  $\beta^{\text{S}}/\beta^{\text{S}}$  “Townes” mice (JAX stock #013071) were lineage-depleted using the lineage cell depletion kit from Miltenyi Biotec. Lin<sup>-</sup> cells were pre-stimulated for 24-hours in StemSpan (Stem Cell Technologies, Vancouver, Canada) supplemented with murine stem cell factor (100ng/mL), human interleukin 11 (100ng/ mL), murine interleukin 3 (20 ng/mL), and human FLT-3 ligand (100 ng/mL). Pre-stimulated Lin<sup>-</sup> cells were then transduced at various MOIs to obtain similar VCNs in the bulk cell product or mock transduced. Twenty-four hours later, one to two million transduced cells were delivered by retro-orbital injection after recipient mice (*Pep Boy/J* [CD45.1 congenic]) were lethally irradiated (1,075 cGy, split in two fractions). A portion of the transduced cells were cultured for 2-week *in vitro* under myeloid differentiation conditions to determine VCN in the cell product.

Peripheral blood samples were collected at weeks 4, 16, and 20 to measure VCN of engrafted cells by ddPCR, expression of Hb $\beta^{\text{AS3}}$  hemoglobin by HPLC, and to determine RBC indices. At week 20, mice were euthanized and BM cells were used to measure engraftment by flow cytometry (CD45.2/CD45.1), VCN, and expression and to seed a methylcellulose based colony forming unit assay. The percentages of sickled erythrocytes present within peripheral blood were also quantified using the Amnis

ImageStream Mark II Imaging Flow Cytometer (Luminex, Austin TX) (as described below).

### **High-performance liquid chromatography**

To characterize and quantify hemoglobin tetramers, including human HbS and HbB<sup>AS3</sup>, and murine HbA and HbF, 1 ul of murine peripheral blood was lysed in 25 ul hemolysate and incubated at room temperature. Hemolysates were then centrifuged at 500g for 10 minutes at 4°C to remove red blood cell ghost. The lysates were then stored frozen at -80°C and later thawed and processed as described by Urbinati et al (Urbinati et al., 2017).

### **CFU progenitor assay**

To quantify VCN range and production of BAS3-globin transcripts in individual BM progenitor cells, total murine BM was acquired from transduced or mock-transduced  $\beta^S/\beta^S$  mice and plated in a methylcellulose medium that supports murine BFU-E growth (Methocult SF M3436; Stem Cell Technologies), as described by Urbinati et al, 2017. Cells were plated at three densities from a serial dilution: 20,000, 60,000 and 120,000 cells per 35 mm gridded plate with two plates per each density. After 14 days of culture at 37°C supplemented with 5% CO<sub>2</sub>, individual BFU-E colonies were enumerated, plucked and separated into two portions for DNA and mRNA isolation (NucleoSpin Tissue XS; Clontech Laboratories, Inc., Mountain View, CA).

### **Cell morphology determination and quantification**

The Amnis ImageStream Mark II Imaging Flow Cytometer was used to determine

and quantify the morphologies of red blood cells isolated from experimental mice at the time of euthanasia. A 2% solution of sodium metabisulfite was made fresh using ultra-distilled water (Invitrogen, Carlsbad, California). Five  $\mu\text{l}$  of blood was mixed with 5  $\mu\text{l}$  of 2% sodium metabisulfite solution (MBS) and incubated for 30 minutes at 37°C. After incubation, 5  $\mu\text{l}$  of the 1:1 blood and MBS solution was resuspended in 95  $\mu\text{l}$  of a 2% glutaraldehyde solution. The 2% glutaraldehyde solution was made by diluting 25% stock glutaraldehyde solution (Sigma-Aldrich, St. Louis, MO) in PBS. Cell images were collected on the ImageStream Flow Cytometer at 60x magnification with bright-field set to channel four and all lasers off. About 40,000 images were collected for each sample and then analyzed using the IDEAS software (Amnis Corporation, Seattle, WA) using a custom template designed with the assistance of Amnis Corporation.

### **Statistical analysis**

All data are reported as mean  $\pm$  standard deviation of the mean unless otherwise stated. Statistical analyses were performed using GraphPad Prism version 7.0 (GraphPad Software, San Diego, California, USA). The statistical significance between two averages was established using unpaired t-test. When the statistical significance between three or more averages were evaluated, a one-way ANOVA was applied followed by multiple paired comparisons for normally distributed data (Tukey test). When normality assumption was violated, Mann-Whitney U test was performed for group-wise comparison instead. Chi-square test were used to compare the frequency of transduced CFUs. Linear regression analyses were used to determine the correlation

between VCN and  $\beta^{\text{AS3}}$ -globin RNA transcripts quantities. All statistical test were two-tailed and a p-value of  $< 0.05$  was deemed significant.

## Tables, Figures and Legends

**Table 1. *In vitro* VCN of gene modified  $\beta^S/\beta^S$  Lin(-) BM Cells before transplant. TU (transduction units); VCN (vector copy number).**

	Lenti/ $\beta$ AS3-FB	Globe-AS3-FB	CoreGA-AS3-FB
<b>Transduction Condition (TU/ml)</b>	<b>6.6x10<sup>6</sup></b>	<b>1x10<sup>7</sup></b>	<b>3.3x10<sup>6</sup></b>
<b><i>In vitro</i> VCN Experiment 1</b>	<b>2</b>	<b>2</b>	<b>2</b>
<b><i>In vitro</i> VCN Experiment 2</b>	<b>2</b>	<b>2</b>	<b>2</b>

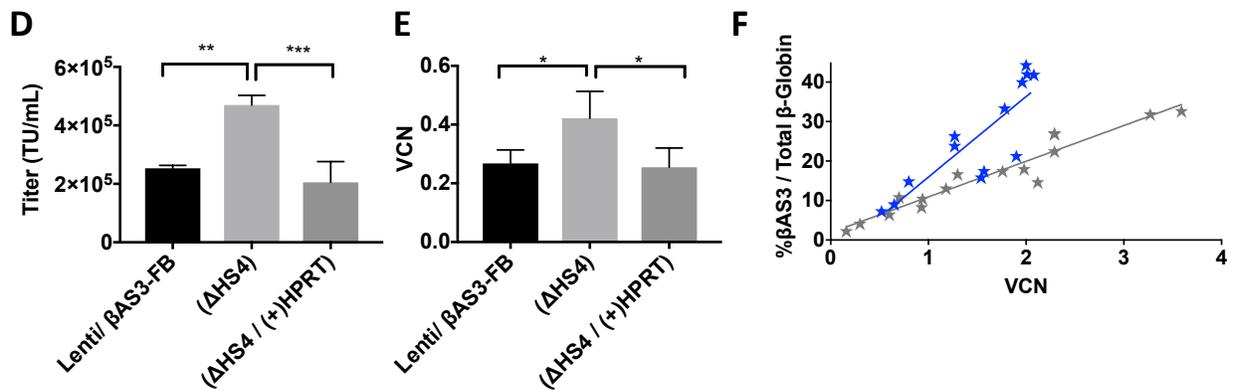
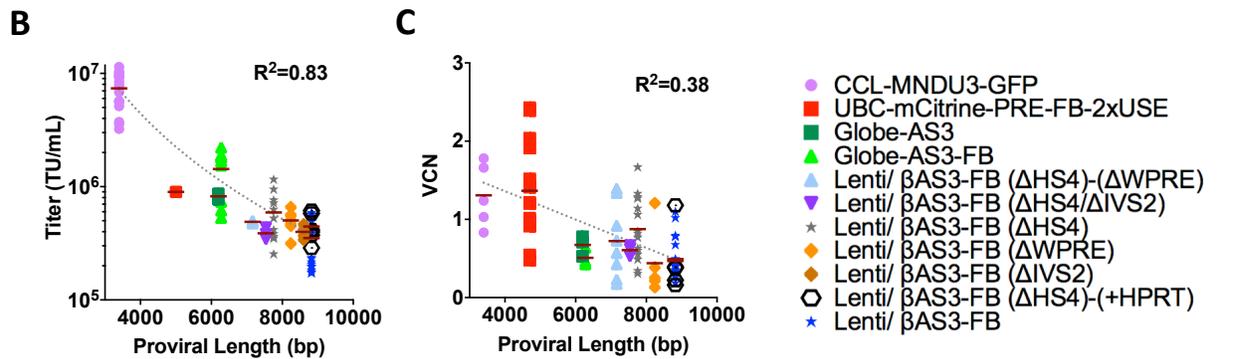
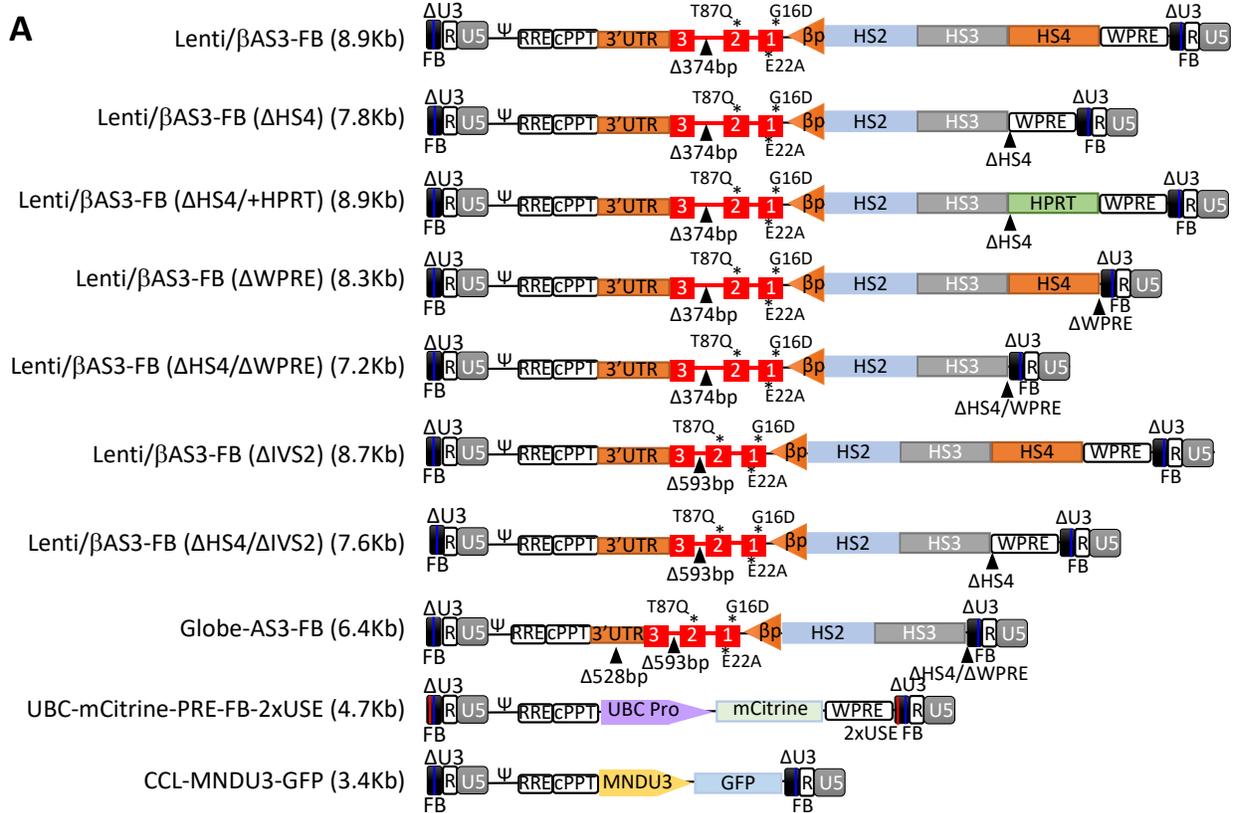
**Table 2. VCNs of bulk bone marrow cells used to seed methylcellulose based CFU assay.**

	<b>Lenti/<math>\beta</math>AS3-FB</b>	<b>Globe-AS3-FB</b>	<b>CoreGA-AS3-FB</b>
<b><i>In vitro</i> CFU Experiment 1</b>	<b>1.6, 1.9</b>	<b>1.9, 1.36</b>	<b>1.9, 2.1</b>
<b><i>In vitro</i> CFU Experiment 2</b>	<b>2.4, 2.6</b>	<b>2.3, 1.9</b>	<b>2.0, 2.6</b>

## Figure 1. Reduction of proviral length increases titer and infectivity

- A) Diagrams of twelve lentiviral vector constructs are shown with their proviral lengths (Kb); the proviral length is the sequence length from start of 5' LTR U3 through end of 3' LTR U5. R (R region of the viral LTR); U5 (U5 region of viral long terminal repeat [LTR]);  $\Psi$ , (packaging signal); RRE (Rev responsive element); cPPT (central polypurine tract); 3'-UTR (3' untranslated region); T87Q, G16D, and E22A (amino acid substitutions);  $\beta$ p, ( $\beta$ -globin promoter); HS2, HS3, HS4 (Locus Control Region [LCR] DNase hypersensitive sites [HS] 2, 3, and 4); WPRE (woodchuck hepatitis virus post-transcriptional regulatory element);  $\Delta$ U3 (self-inactivating deletion in the U3 region); FB (FII-BEAD insulator); All major alterations introduced into each derivative are emphasized by triangle.
- B) Constructs were packaged and titered and the quantity of infectious particles were plotted as a function of proviral length (bp). Each point in the plot represents an individual 10-cm plate of virus packaged and titered in parallel with CCL-MNDU3-GFP, Lenti/ $\beta$ AS3-FB, Globe-AS3-FB and various derivatives.
- C) Human CD34+ hematopoietic stem and progenitor cells (HSPCs) were transduced with constructs at  $2 \times 10^7$  TU/mL and cultured under myeloid differentiation conditions to assess levels of infectivity. Each point in the plot represents an individual transduction, compared in parallel with CCL-MND-GFP (or CCL-UBC-mCitrine-PRE-FB-2xUSE), Lenti/ $\beta$ AS3-FB, Globe-AS3-FB, and various derivatives.
- D) Quantity of infectious particles of constructs when packaged and titered in parallel. n=6. \*\*p<0.01; \*\*\*p<0.001.

- E) Vector copy number (VCN) was determined by digital droplet polymerase chain reaction (ddPCR) 14 days after Human CD34+ HSPCs were transduced at  $2 \times 10^7$  TU/mL and cultured under myeloid differentiation conditions to assess infectivity.  $n=3$ .  $*p < 0.05$ .
- F) Percent  $\beta^{AS3}$ -globin RNA expression was determined by reverse transcription ddPCR and shown in relation to VCN for Lenti/ $\beta$ AS3-FB (blue [ $n=13$ ]) and Lenti/ $\beta$ AS3-FB-( $\Delta$ H54) (gray [ $n=16$ ]). Each point represents an individual transduction done in parallel with Lenti/ $\beta$ AS3-FB. Difference between slopes  $p=0.0010$ .



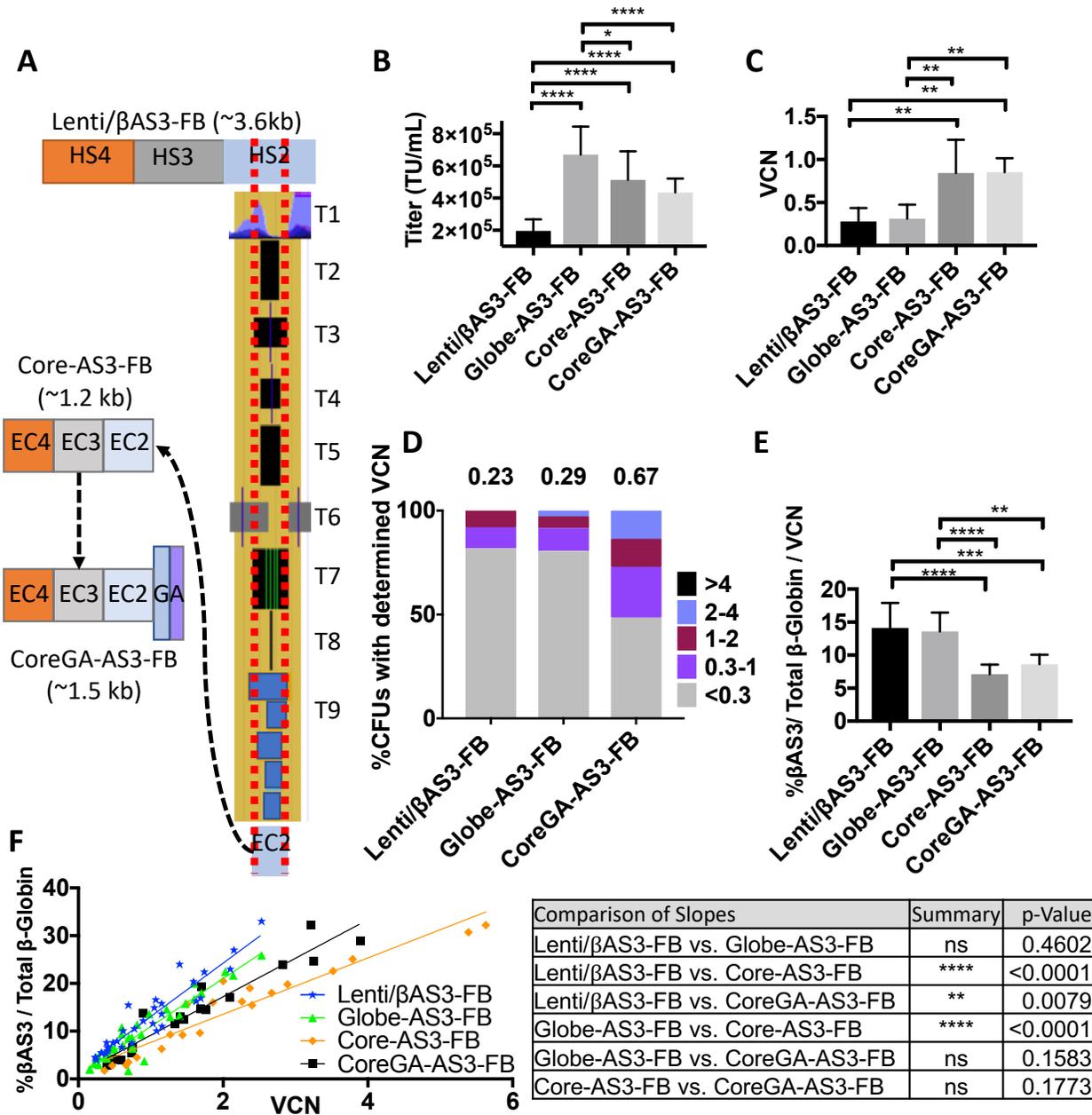
**Figure 2. Approach for length reduction and *In vitro* characterization of reduced length vectors**

- A) Published genomic data available through ENCODE were used to redefine the putative HS boundaries. The track sets (T) 1-8 were used to generate new boundaries for the LCR's HS sequences. T1 (H3k27ac); T2 (DNase I [Encode]); T3 (DNase I [University Wisconsin]); T4 (DNase I [UNC Chapel Hill]); T5 (Formaldehyde-Assisted Isolation of Regulatory Elements:FAIRE [UNC Chapel Hill]); T6 (DNase I/FAIRE/ChiP-Seq [Duke/UNC Chapel Hill/UT-Austin]); T7 (Transcription factor ChiP-Seq); T8 (conserved TF motifs); T9 (Regulatory elements [Open Regulatory Annotation database]). These new sequences were termed "ENCODE Cores" (EC) 2, 3, and 4 and combined to create Core  $\beta$ -LV. The 143bp human Ankyrin Element was then combined with the 218bp murine Gata1 HS4 element and added to EC2, 3 and 4 to create CoreGA-AS3-FB  $\beta$ -LV.
- B) Quantity of infectious particles of constructs when packaged and titered in parallel. n=5. \*p<0.05; \*\*\*\*p<0.0001
- C) Human CD34+ HSPCs were transduced with constructs at  $1 \times 10^7$  TU/mL and cultured under myeloid culture conditions for 14-days. VCN was determined by ddPCR. n=3. \*\*p<0.01
- D) Human CD34+ HSPCs were transduced in parallel at  $1 \times 10^6$  TU/mL and used to seed a methylcellulose-based colony forming unit assay. Shown are stacked bar-graphs representing percentages of colonies with a determined VCN range. n=50

colonies, Lenti/ $\beta$ AS3-FB; n=36 colonies, Globe-AS3-FB; n=37 colonies, CoreGA-AS3-FB.

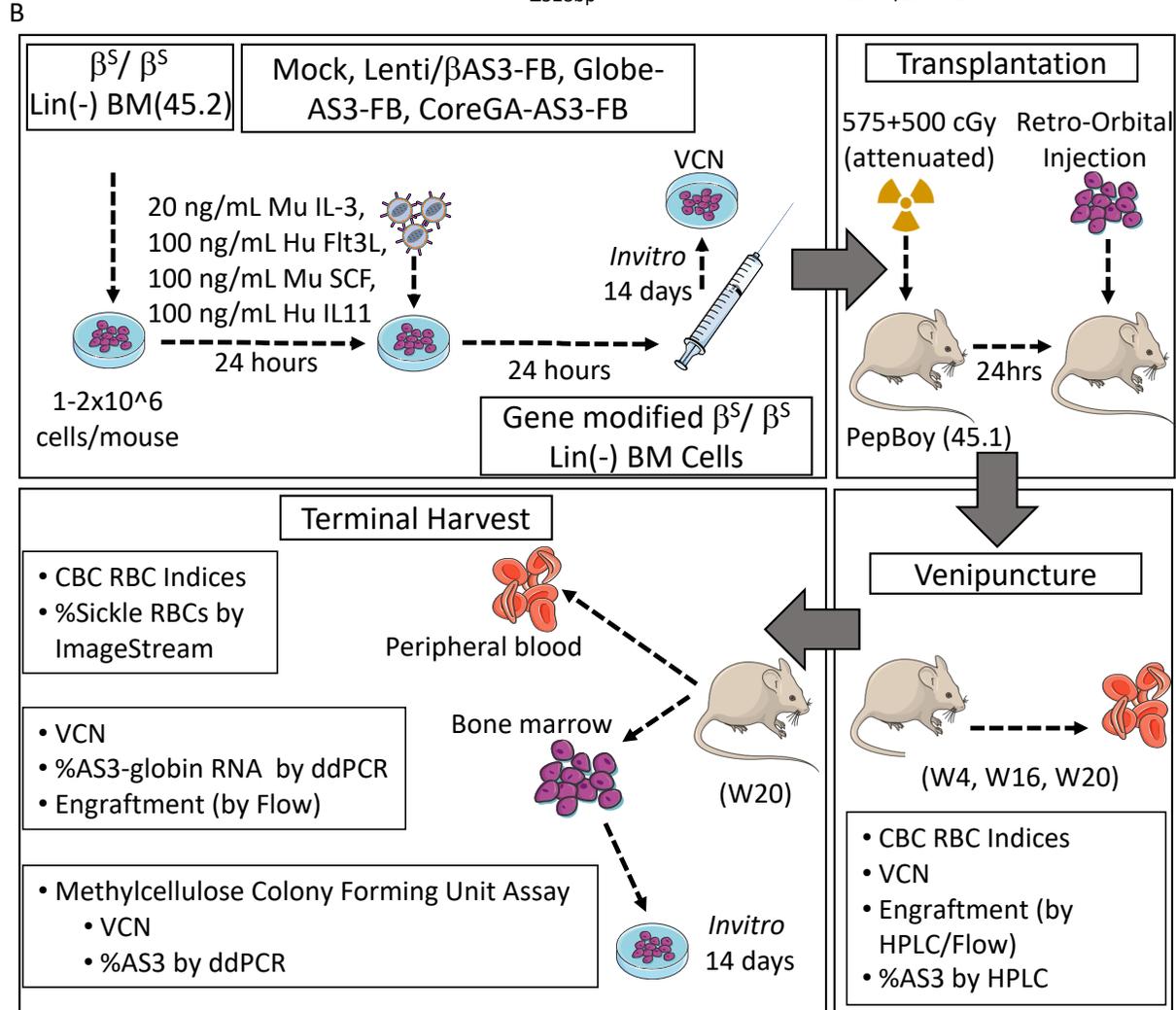
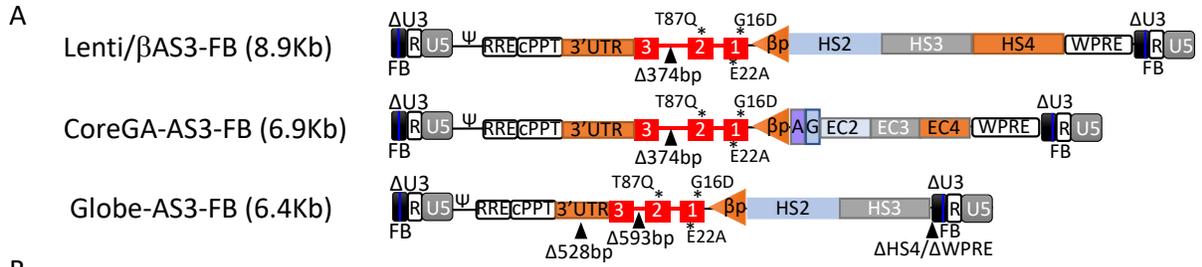
E) Human CD34+ HSPCs were transduced in parallel at  $1 \times 10^7$  TU/mL and cultured under erythroid culture conditions for 14-days. Expression of different constructs is presented as percentage of  $\beta$ AS3-globin normalized to VCN. n=9. \*\*\*p<0.001

F) Human CD34+ HSPCs were transduced at various multiplicities of infection (MOI) and differentiated under erythroid culture conditions for 14-days. Percent  $\beta$ AS3-globin RNA expression was determined by reverse transcription ddPCR and is shown in relation to VCN. Each point in the plot represents an individual transduction done in parallel with Lenti/ $\beta$ AS3-FB and Globe-AS3-FB. n=33, Lenti/ $\beta$ AS3-FB; n=36, Globe-AS3-FB; n=21, Core; n=18, CoreGA-AS3-FB.



**Figure 3. Outline of in vivo experiments in a sickle cell disease (SCD) mouse model**

- A) Diagrams of three lentiviral vectors assayed in vivo.
- B) To characterize vector performance in long-term hematopoietic repopulating cells, lineage-depleted (Lin (-)) bone marrow (BM) cells derived from homozygous  $\beta^S/\beta^S$  mice were transduced at different MOIs to obtain similar VCNs in the bulk cell product or mock transduced. Cells were retained for in vitro culture to determine VCN by ddPCR. One to two million transduced cells were delivered by retro-orbital injection after recipient mice (Pep Boy/J [CD45.1 congenic]) were lethally irradiated. Peripheral blood was collected at weeks 4 (W4), 16 (W16), and 20 (W20) to measure engraftment, transduction efficiency, expression, and RBC indices. At week 20 mice were euthanized and BM cells were used to measure engraftment, transduction efficiency, and expression and to seed a methylcellulose-based colony forming unit assay. The percentages of sickled erythrocytes present within peripheral blood were also quantified at week 20. Mu IL-3 = murine Interleukin-3; Hu Flt3L = human Flt-3 ligand; Mu SCF = murine stem cell factor; Hu IL11= human interleukin-11.



**Figure 4. In vivo analysis of hematologic correction of SCD mouse model.** Mice were bled at weeks 4, 16, and 20 and peripheral blood (PB) analyzed. Only mice with >97% BM donor engraftment were included in analysis. n=9, Mock; n=8, Lenti/ $\beta$ AS3-FB; n=9, Globe-AS3-FB; n=8, CoreGA-AS3-FB.

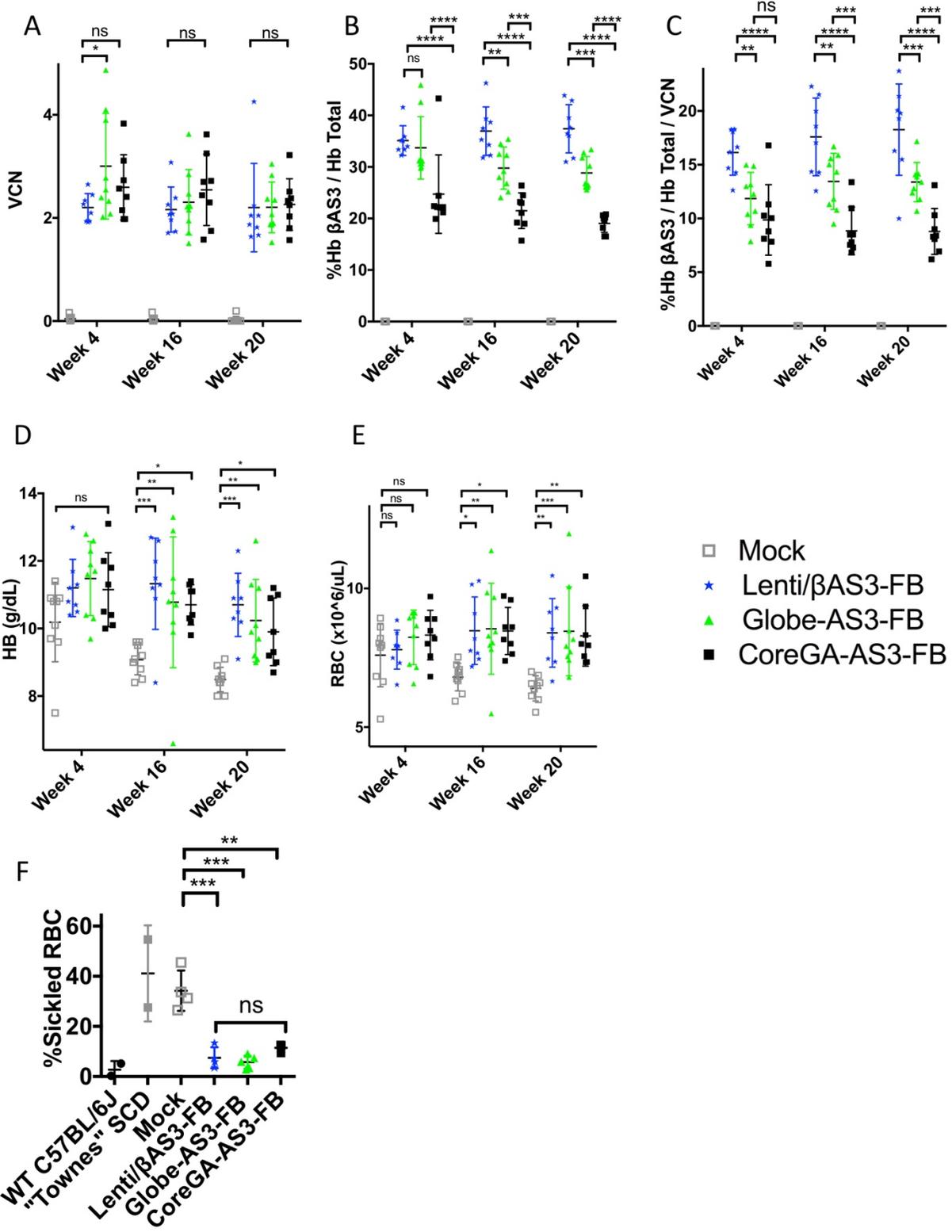
A) Peripheral blood VCN by ddPCR. \*\*\*\*p<0.0001

B) Percentages of Hb  $\beta$ AS3 tetramers in PB lysates measured by high-performance liquid chromatography. \*p<0.05

C) Percentages of Hb  $\beta$ AS3 tetramers normalized to PB VCN.

D) Hemoglobin (HB [g/dL]) levels. \*\*p<0.01

E) Red Blood Cell (RBC) count (x10<sup>6</sup>). \*\*\*p<0.001

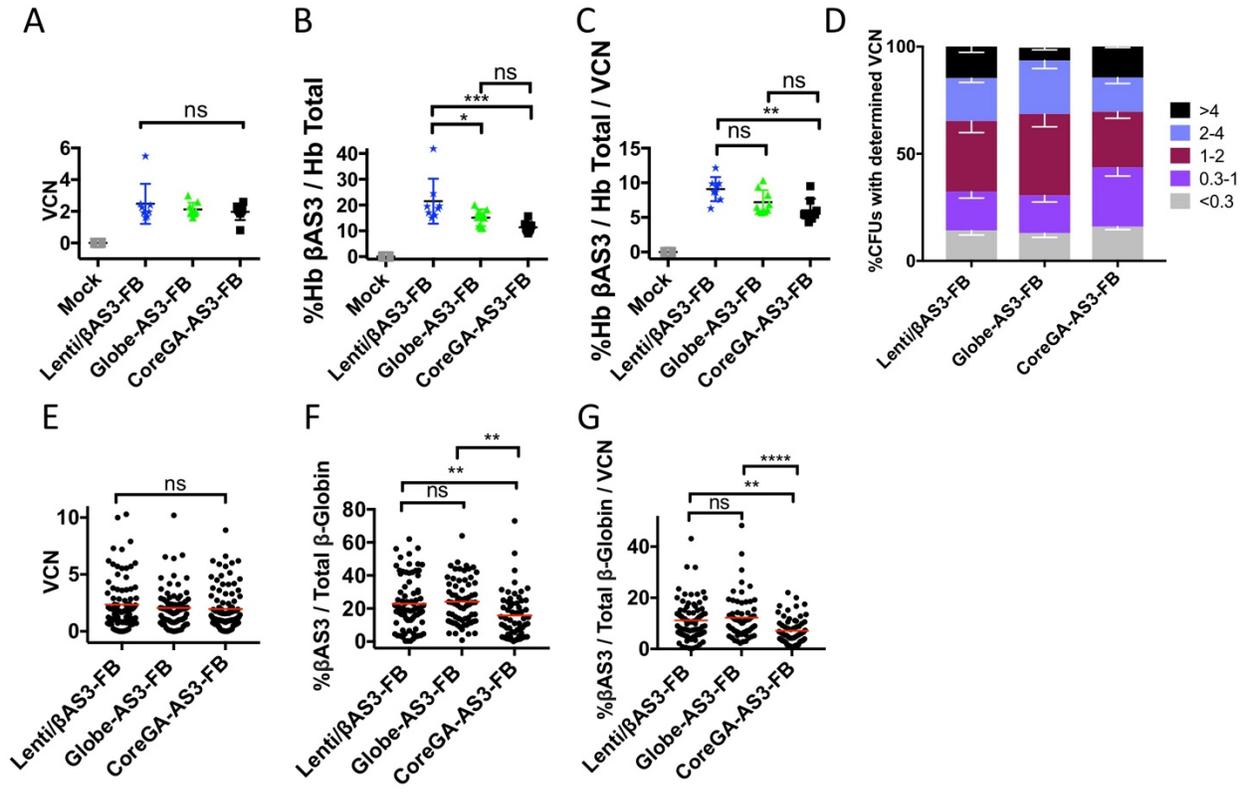


**Figure 5. In vivo analysis of bone marrow  $\beta$ AS3-globin expression from SCD**

**mouse model.** Whole BM was taken from each mouse at time of euthanasia and mice with >97% donor engraftment were analyzed. n=9, Mock; n=8, Lenti/ $\beta$ AS3-FB; n=9, Globe-AS3-FB; n=8, CoreGA-AS3-FB.

- A) BM VCN by ddPCR. ns= not significant.
- B) Percentages  $\beta$ AS3-globin RNA expression of BM determined by reverse transcription ddPCR. \*p<0.05; \*\*p<0.01
- C) Percentage of BM  $\beta$ AS3-globin RNA expression normalized to BM VCN.
- D) BM obtained at time of euthanasia was used to seed a methylcellulose based CFU assay. Individual colonies were plucked two weeks after plating and gDNA and RNA extracted. Shown are stacked bar-graphs representing percentages of colonies with a determined VCN range. Shown are mean (+/-) SEM from two independent experiments (n = 2 donors / arm). n=189 colonies, Lenti/ $\beta$ AS3-FB; n=129 colonies, Globe-AS3-FB; n=156 colonies, CoreGA-AS3-FB.
- E) VCN of CFUs from two independent experiments (n = 2 donors / arm). Same data as (D). Only PCR-positive CFUs with detectable endogenous globin expression were included. n=93 colonies, Lenti/ $\beta$ AS3-FB; n=84 colonies, Globe-AS3-FB; n=94 colonies, CoreGA-AS3-FB. Each dot represents a single colony. Bar represents grand average of VCN from two independent experiments.
- F) Percent  $\beta$ AS3-globin RNA expression was determined by reverse transcription ddPCR. \*\*\*\*p<0.0001
- G) Percent  $\beta$ AS3-globin RNA expression normalized to VCN for matched colony. \*\*\*p<0.001.

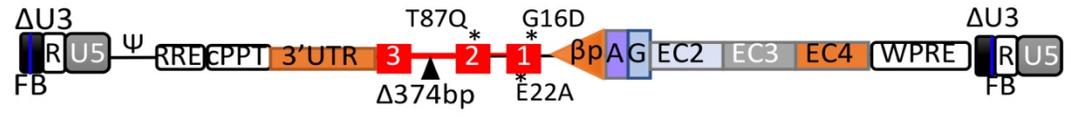
H) Percentage of sickled red blood cells in PB quantified by ImageStream at 20 weeks post transplantation. “Townes” SCD (homozygous SCD mice), included as negative control; WT C57BL/6J (Pep Boy/J [CD45.1 congenic]) mice are included as a positive control. Only mice from experiment two with >97% BM donor engraftment were included in analysis. n=2, WT C57BL/6J; n=1, “Townes” SCD; n=4, Mock; n=4, Lenti/ $\beta$ AS3-FB; n=5, Globe-AS3-FB; n=3, CoreGA-AS3-FB.



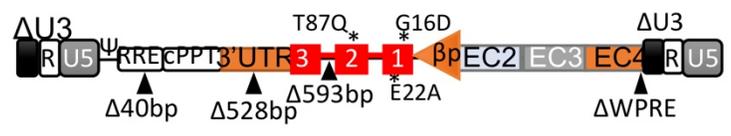
## Figure 6. Further length reduction offers improved vector performance

- A) Diagrams of lentiviral vector constructs. All annotations similar to those shown in figure (3A).
- B) Quantity of infectious particles of constructs when packaged and titered in parallel. n=3. \*\*\*\*p<0.0001
- C) Human CD34+ HSPCs were at various MOIs and cultured under myeloid culture conditions for 14-days. VCN was determined by ddPCR. n=3. \*p<0.05; \*\*p<0.01
- D) Human CD34+ HSPCs were transduced in parallel at  $1 \times 10^7$  TU/mL and cultured under erythroid culture conditions for 14-days. Expression of different constructs is presented as percentage of  $\beta$ AS3-globin normalized to VCN. n=9.

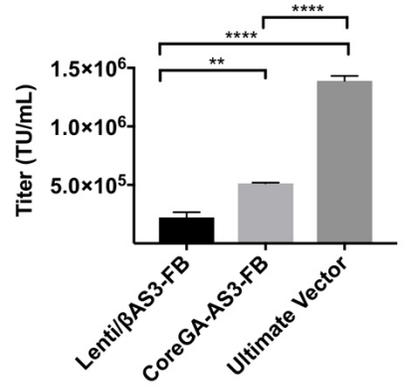
**A**  
CoreGA-AS3-FB (6.9Kb)



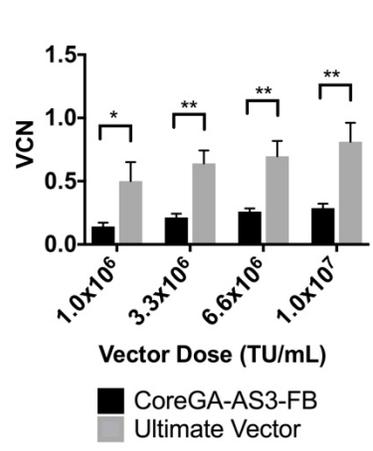
Ultimate Vector (4.7Kb)



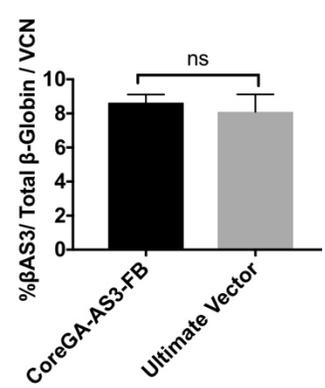
**B**



**C**



**D**



## Supplemental results:

### Vector performance can be improved by removal of additional sequence elements

Data generated from dissecting the differences between Lenti/ $\beta$ AS3-FB and Globe-AS3-FB indicated that Lenti/ $\beta$ AS3-FB's titer and infectivity could be improved through reducing vector length. We hypothesized that a decrease of 2-3Kb would be required to generate a construct produced at three- to four-fold higher raw titer with superior gene transfer to HSPCs when compared to Lenti/ $\beta$ AS3-FB at equivalent MOI. To realize a reduced length vector, we redefined the boundaries of the LCR HS elements and added elements to facilitate position independent expression of  $\beta^{\text{AS3}}$ -globin, reducing overall proviral length from ~8.9kb to ~6.8kb.

The dissection of differences between Lenti/ $\beta$ AS3-FB and Globe-AS3-FB revealed that vector performance could be further improved through removal of the ~600bp WPRE. Although inclusion of WPREs into LVs have been shown to increase titer and transgene expression (Hope, 2002), Lenti/ $\beta$ AS3-FB would not benefit from increased  $\beta^{\text{AS3}}$ -globin expression conferred by the WPRE upon due to the reversed orientation of the human  $\beta$ -globin cassette within the vector and the positioning of the WPRE at the other end of the vector between the HS4 element and 3' long terminal repeat. The WPRE may act to increase vector titer by increasing the amount of genomic RNA in packaging cells.

Paradoxically, removal of WPRE resulted in increased titer ( $p < 0.01$ ) (Supplemental Figure 1A). While there was a trend toward decreased gene transfer to CD34+ cells at a fixed MOI conferred upon the deletion of the WPRE, the decrease in gene transfer did not reach statistical significance (Supplemental Figure 1B). These

finding on the effects of WPRE removal were additionally replicated in the HS4-deleted Lenti/ $\beta$ AS3-FB derivatives (Supplemental Figure 1C and 1D). Potentially, the benefits of WPRE on increasing vector transcripts and hence titer are counter-acted by its adverse effect on overall vector genome size in these already large vectors.

The intervening sequence 2 (IVS2) deletions present in Lenti/ $\beta$ AS3-FB (372 bp) and Globe-AS3-FB (652 bp) were shown to increase titer and provide stable proviral transmission of  $\beta$ -LVs to primary cells through removing repeat sequences deemed detrimental to reverse transcription (“A normal level of beta-globin expression in erythroid cells after retroviral cells transfer.,” 1992; “Efficient 3'-end formation of human beta-globin mRNA in vivo requires sequences within the last intron but occurs independently of the splicing reaction.,” 1998; “Mutagenesis of retroviral vectors transducing human beta-globin gene and beta-globin locus control region derivatives results in stable transmission of an active transcriptional structure.,” 1994; “The human beta-globin gene contains multiple regulatory regions: identification of one promoter and two downstream enhancers.,” 1988; “Two 3' sequences direct adult erythroid-specific expression of human beta-globin genes in transgenic mice.,” 1987). When we compared a vector with a larger 652bp IVS2 deletion [Lenti/ $\beta$ AS3-FB ( $\Delta$ IVS2)] head-to-head against Lenti/ $\beta$ AS3-FB (a construct containing the smaller 372bp IVS2 deletion), no differences in titers were observed (Supplemental Figure 2A). Surprisingly, however, a 1.4-fold increase in normalized expression was observed for Lenti/ $\beta$ AS3-FB ( $\Delta$ IVS2) ( $p < 0.05$ ) (Supplemental Figure 2B). This finding suggested that incorporation of the larger IVS2 deletion could compensate for the decrement of expression incurred by deletion of HS4. To confirm that introduction of the larger IVS2 deletion could increase

expression per vector genome, the larger IVS2 deletion was introduced into a construct lacking the HS4 element to produce Lenti/ $\beta$ AS3-FB ( $\Delta$ HS4/ $\Delta$ IVS2) and compared head-to-head against Lenti/ $\beta$ AS3-FB ( $\Delta$ HS4). While no differences in titers were observed upon introduction of the larger IVS2 deletion (Supplemental Figure 2C), expression per VCN for the construct containing both the HS4 and larger IVS2 deletions were surprisingly comparable to the parent Lenti/ $\beta$ AS3-FB vector (Supplemental Figure 2D). This finding confirmed that the larger IVS2 deletion benefited expression. Taken together, these data demonstrated that an additional ~1 kb of sequence could be removed from Lenti/ $\beta$ AS3-FB (removal of WPRE and introduction of larger IVS2 deletion) without loss to titer, transduction efficiency or  $\beta^{\text{AS3}}$ -globin expression.

### **Expression per vector genome can be improved**

We explored several different strategies to improve expression per vector genome, while retaining the smallest proviral DNA footprint possible, as increased expression was deemed potentially beneficial in clinical situations of poor gene transfer.

A Lenti/ $\beta$ AS3-FB derivative was designed with an extended promoter of 615bp (Lenti/ $\beta$ AS3-FB [LongPro]) (vs. 265 bp in the standard vectors) (Supplemental Figure 3A). To determine if this longer  $\beta$ -globin promoter could increase transgene expression per VCN, HSPCs were transduced with Lenti/ $\beta$ AS3-FB and its derivative at equal MOI and cultured under erythroid culture conditions for 14 days. After five independent experiments, a 1.2-fold increase in expression per vector genome was observed when comparing LongPro to its parental construct ( $p < 0.05$ ) (Supplemental Figure 3B).

Previous reports have demonstrated that the full length HS1 element could increase expression per vector genome when combined with the full length HS 2,3, and 4 elements(Lisowski and Sadelain, 2007). We hypothesized that the HS1 ENCODE core (EC1) could similarly improve expression. Thus, as an alternative strategy to increase normalized expression, we again deployed ENCODE to redefine the boundaries of the HS1 element and designed a derivative of Core-AS3-FB containing the EC1 and EC2,3 and 4 ((called +EC1) to determine the influence of EC1 on  $\beta^{\text{AS3}}$ -globin expression. Healthy donor HSPCs were transduced with Core-AS3-FB and +EC1 at equal MOI and cultured under erythroid culture conditions for 14 days. From three independent experiments, a 1.6-fold increase in expression per vector genome was observed when comparing Core-AS3-FB(+EC1) to its parental construct ( $p < 0.05$ ) (Supplemental Figure 3C).

Finally, we tested the usefulness of codon optimization on  $\beta^{\text{AS3}}$ -globin transgene expression. The  $\beta^{\text{AS3}}$ -globin transgene was codon optimized for expression using the J-CAT software(Grote et al., 2005) and care was taken to retain the three anti-sickling mutations. The optimized  $\beta^{\text{AS3}}$ -globin exons had 84% sequence homology at the nucleotide level when compared to the original unmodified sequences. Three derivatives of Core-AS3-FB were made that contained, 1). Codon-optimized exon one with unmodified exons two and three, 2), codon-optimized exons one and two with unmodified exon three, and 3), codon-optimized exons one, two, and three with no unmodified exons. To determine if codon optimization could enhance expression, HSPCs derived from multiple healthy donors were transduced with Core-AS3-FB derivatives at equal MOI and subjected to *in vitro* erythroid differentiation for 14 days.

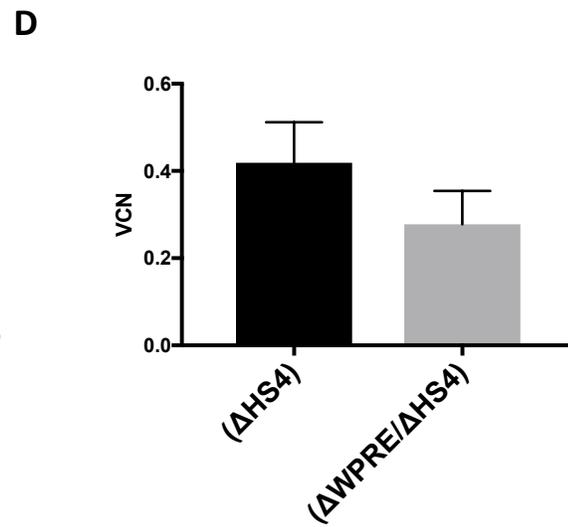
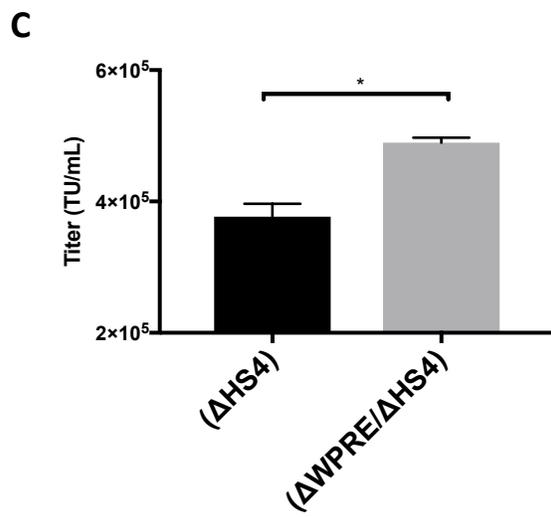
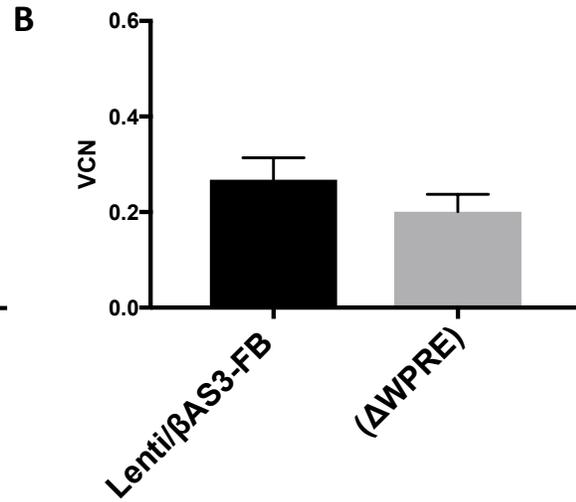
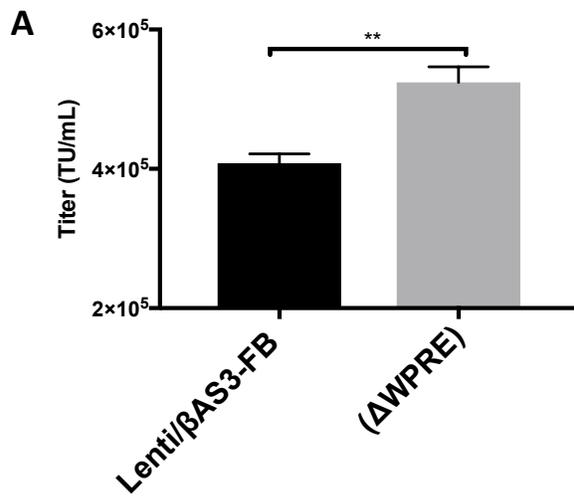
Enhancement of expression by about 1.6-fold was observed in derivatives that contained a codon optimized exon two ( $p < 0.05$ ) when compared to the parental construct. The additional optimization of exons one or three did not result in increased expression per vector genome when compared to the appropriate control (Supplemental Figure 3D).

Taken together, these data provide strategies to increase expression per vector genome. Further exploration of core vector derivatives will identify the best configuration of elements in a  $\beta$ -globin vector of reduced length.

### **Supplemental Figure 1. Assessment of requirement for WPRE**

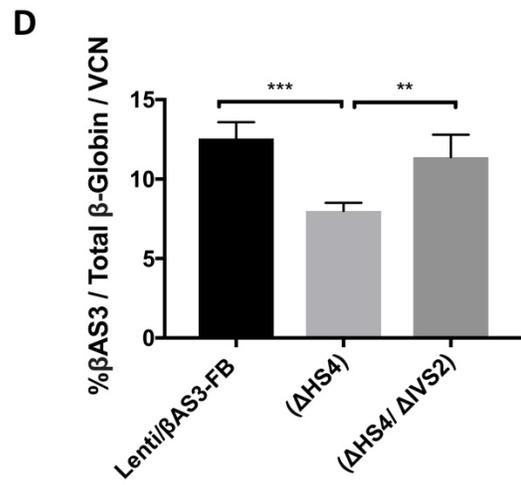
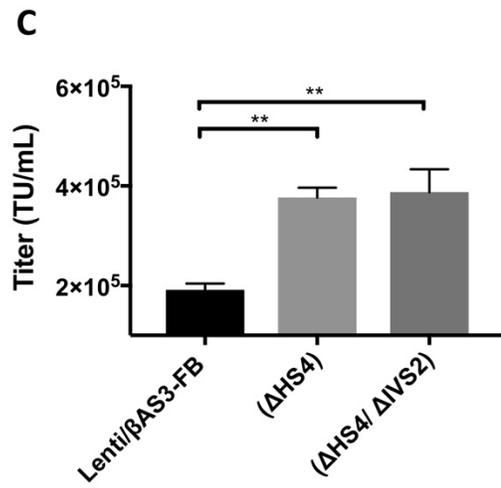
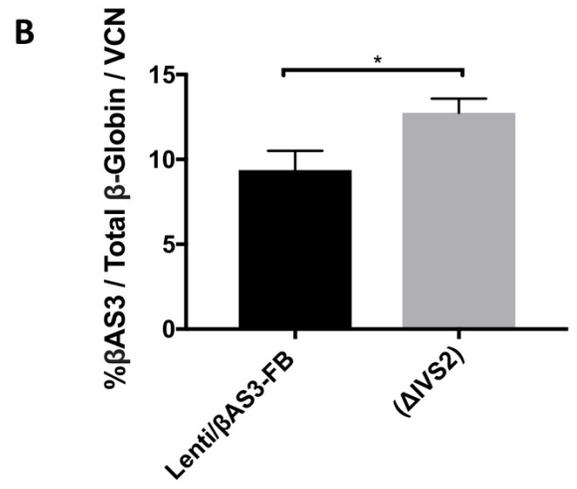
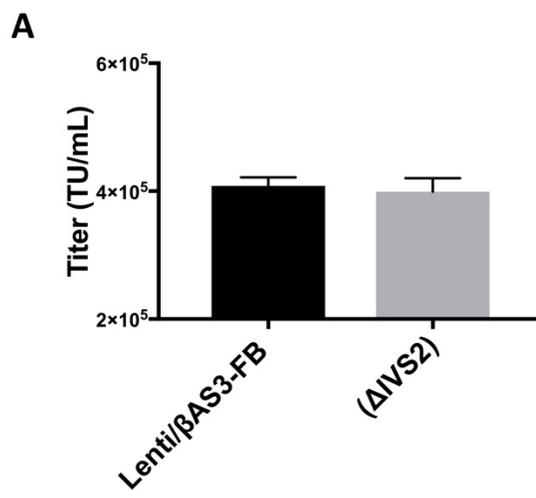
A) and C) Quantity of infectious particles of constructs [parent vector Lenti/ $\beta$ AS3-FB, and derivatives Lenti/ $\beta$ AS3-FB ( $\Delta$ WPRE) (abbreviated  $\Delta$ WPRE), and Lenti/ $\beta$ AS3-FB ( $\Delta$ WPRE/ $\Delta$ HS4) (abbreviated  $\Delta$ WPRE/ $\Delta$ HS4)] when packaged and titered in parallel. n=3. \*p<0.05; \*\*p<0.01

B) and D) Human CD34+ HSPCs were transduced with constructs at  $1 \times 10^7$  TU/mL and cultured under myeloid culture conditions for 14 days. VCN was determined by ddPCR. n=3.



**Supplemental Figure 2. Evaluation of larger IVS2 sequence deletion,**

- A) and C) Quantity of infectious particles of constructs [parent vector Lenti/ $\beta$ AS3-FB, and derivatives Lenti/ $\beta$ AS3-FB ( $\Delta$ HS4) (abbreviated  $\Delta$ HS4), and Lenti/ $\beta$ AS3-FB ( $\Delta$ HS4/ $\Delta$ IVS2) (abbreviated  $\Delta$ HS4/ $\Delta$ IVS2)] when packaged and titered in parallel. n=3. \*\*p<0.01.
- B) and D) Human CD34+ HSPCs were transduced in parallel at  $2 \times 10^7$  TU/mL and cultured under erythroid culture conditions for 14 days. Expression of different constructs is presented as percentage of  $\beta^{AS3}$  normalized to VCN. n=5. \*p<0.05; \*\*\*p<0.001



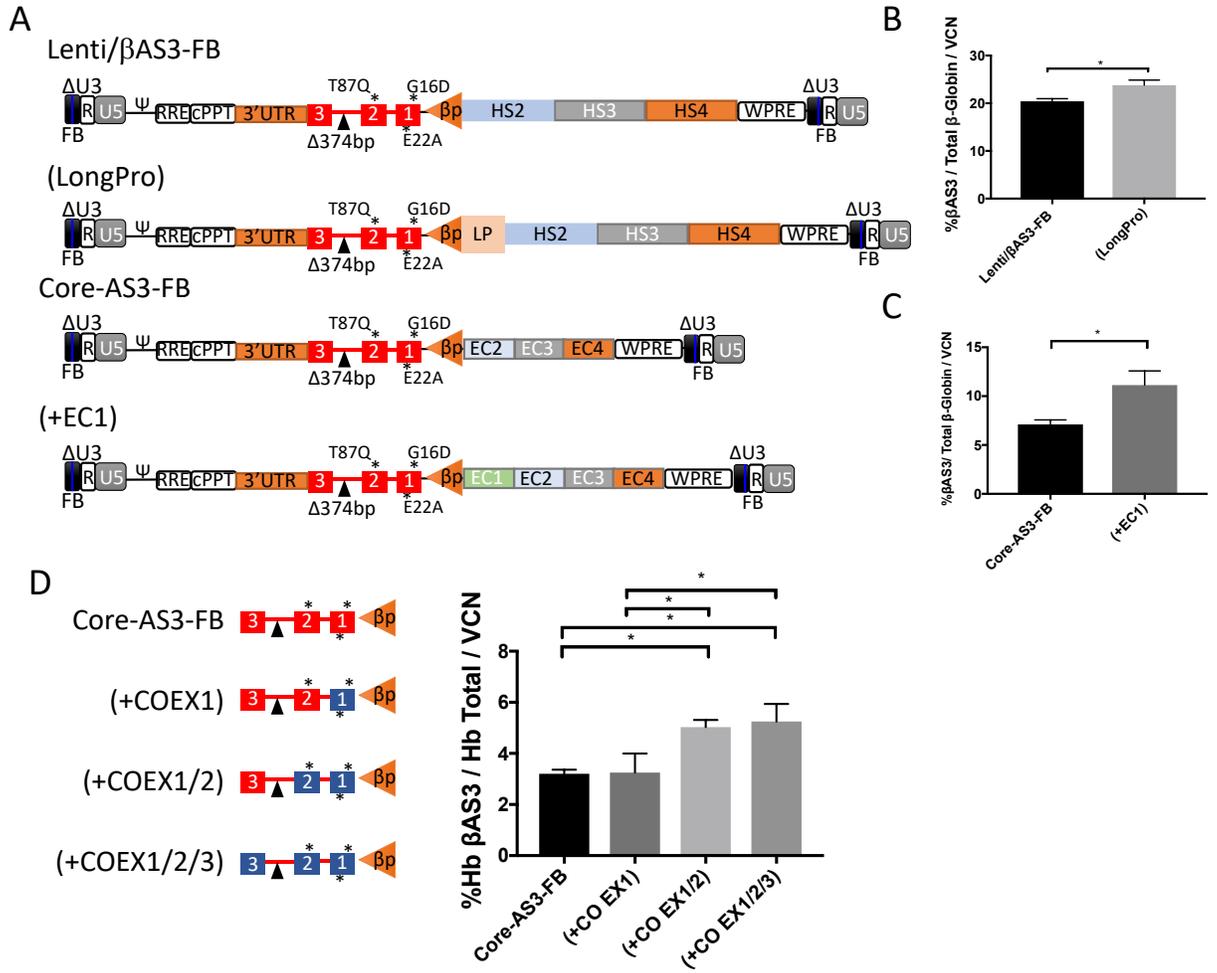
**Supplementary Figure 3. Modifications increase expression per vector genome**

A) Diagrams of lentiviral vector constructs. LP (Longer  $\beta$ -globin promoter); EC1 (ENCODE core 1); all other annotations similar to those shown in figure (3A).

B) Human CD34+ HSPCs were transduced in parallel at  $2 \times 10^7$  TU/mL and cultured under erythroid culture conditions for 14-days. Expression of different constructs is presented as percentage of  $\beta$ AS3-globin normalized to VCN. n=5. \*p<0.05

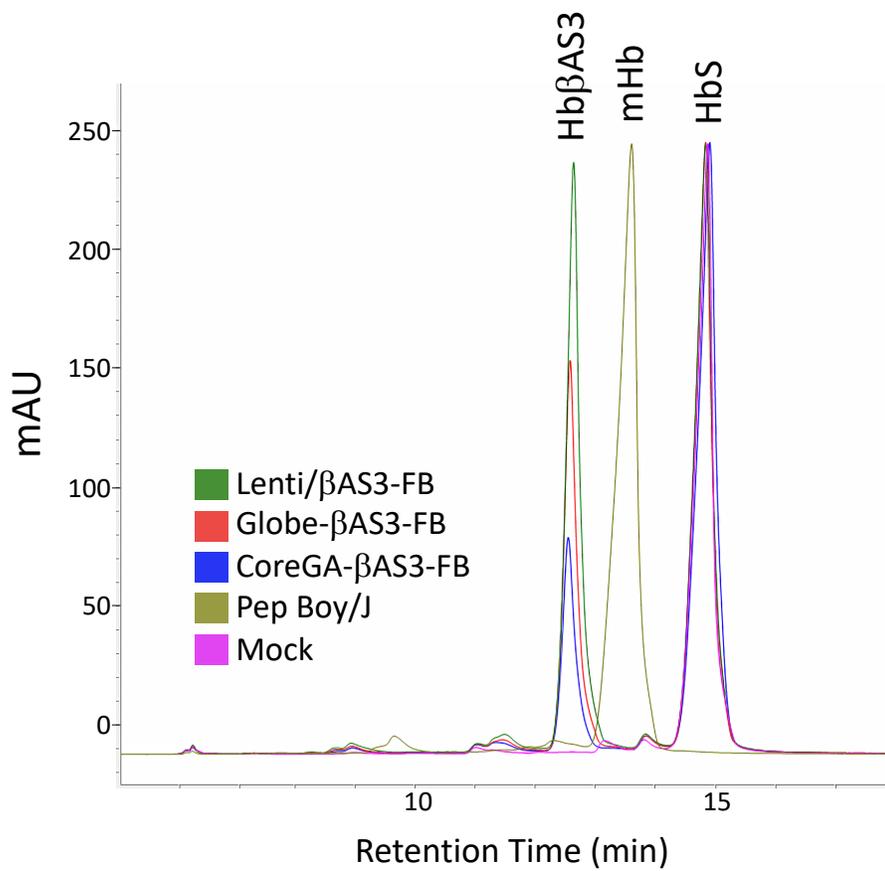
C) Human CD34+ HSPCs were transduced in parallel at  $1 \times 10^7$  TU/mL and cultured under erythroid culture conditions for 14-days. Presentation of results the same as (B). n=3.

D) Diagrams showing modified exons (in blue) of  $\beta$ AS3 transgene. All modifications were assessed in Core-AS3-FB backbone. Human CD34+ HSPCs were transduced in parallel at  $2 \times 10^7$  TU/mL and cultured under erythroid culture conditions for 21-days. Expression of different constructs is presented as percentages of Hb  $\beta$ AS3 tetramers from cell lysates normalized VCN. n=3.



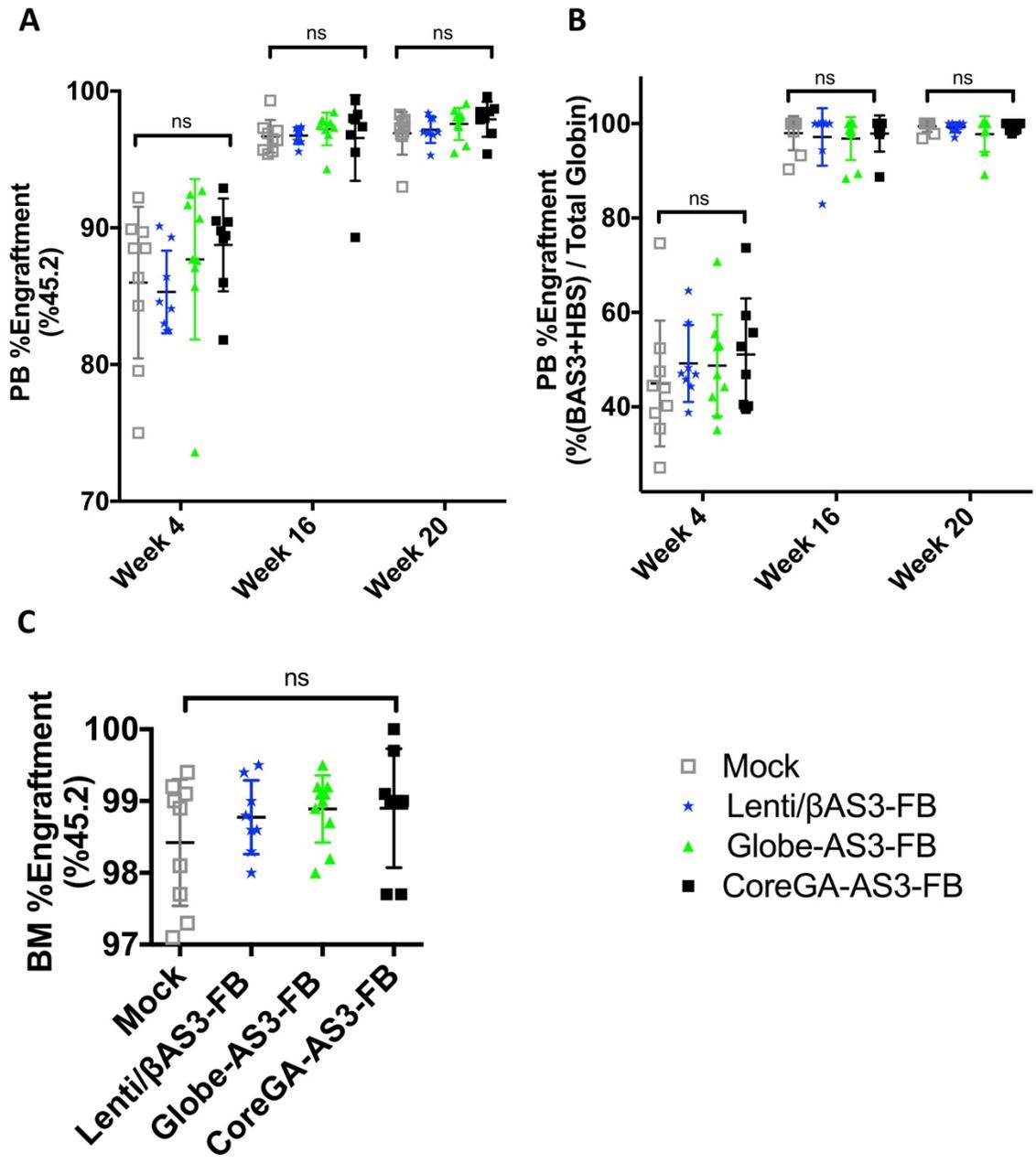
**Supplementary Figure 4. High-performance liquid chromatography (HPLC) chromatograms from peripheral blood lysates obtained from representative mice of each experimental arm. Hb $\beta^{AS3}$ , HbS, and mHB (murine HB) peaks are visible. mAU, milli-Absorbance Units; min, minutes.**

A



**Supplementary Figure 5. Peripheral blood and bone marrow engraftment levels**

- A. Peripheral blood engraftment levels at weeks 4, 16, and 20 quantified by FACS.
- B. Peripheral blood engraftment levels at weeks 4, 16, and 20 quantified by HPLC.
- C. Bone marrow engraftment levels at week 20 quantified by FACS.



## References

1. Gene mutations in human haemoglobin: the chemical difference between normal and sickle cell haemoglobin. 180, 326–328 (1957).
2. Francis, R. J. & Johnson, C. S. Vascular occlusion in sickle cell disease: current concepts and unanswered questions. 77, 1405–1414 (1991).
3. Effect of Hydroxyurea on the Frequency of Painful Crises in Sickle Cell Anemia. 332, 1317–1322 (1995).
4. Niihara, Y. et al. A Phase 3 Trial of L-Glutamine in Sickle Cell Disease. 379, 226–235 (2018).
5. Albuerne, E. P. Allogeneic donor availability for hematopoietic stem cell transplantation in children with sickle cell disease. 62, 1285–1287 (2015).
6. Urbinati, F. et al. Preclinical studies for a phase 1 clinical trial of autologous hematopoietic stem cell gene therapy for sickle cell disease. 19, 1096–1112 (2017).
7. Ribeil, J.-A. et al. Gene Therapy in a Patient with Sickle Cell Disease. 376, 848–855 (2017).
8. Cavazzana-Calvo, M. & Kalle, von, C. LentiGlobin gene therapy for transfusion-dependent  $\beta$ -Thalassemia: Update from the northstar HGB-204 phase 1/2 clinical study. 19, S19 (2017).
9. Kanter, J. et al. Interim results from a phase 1/2 clinical study of lentiglobin gene therapy for severe sickle cell disease. (2016).
10. Walters, M. C., Thrasher, A. J. & Yannaki, E. Results from the Hgb-207 (Northstar-2) trial: A phase 3 study to evaluate safety and efficacy of lentiglobin gene

therapy for transfusion-dependent  $\beta$ -thalassemia (TDT) in patients with non- $\beta^0/\beta^0$  genotypes. (2017).

11. Thompson, A. A. et al. Gene Therapy in Patients with Transfusion-Dependent beta-Thalassemia. *N. Engl. J. Med.* 378, 1479–1493 (2018).

12. Cooper, A. R. et al. Highly efficient large-scale lentiviral vector concentration by tandem tangential flow filtration. *177*, 1–9 (2011).

13. Cavazzana-Calvo, M. et al. Transfusion independence and *HMGA2* activation after gene therapy of human  $\beta$ -thalassaemia. *Nature* 467, 318–322 (2010).

14. Beschorner, N. et al. Improving Lentiviral Transduction of CD34 +Hematopoietic Stem and Progenitor Cells. *29*, 104–113 (2018).

15. Anastasov, N., Höfig, I., Mall, S., Krackhardt, A. M. & Thirion, C. in *Lentiviral Vectors and Exosomes as Gene and Protein Delivery Tools* 1448, 49–61 (Humana Press, New York, NY, 2016).

16. Cavazzana, M. & Six, E. A Nontoxic Transduction Enhancer Enables Highly Efficient Lentiviral Transduction of Primary Murine T Cells and Hematopoietic Stem Cells. *10*, 341–347 (2018).

17. Petrillo, C. et al. Cyclosporine H Overcomes Innate Immune Restrictions to Improve Lentiviral Transduction and Gene Editing In Human Hematopoietic Stem Cells. *Cell Stem Cell* 23, 820–832.e9 (2018).

18. Sanvito, F. In vivo selection of genetically modified erythroblastic progenitors leads to long-term correction of  $\beta$ -thalassemia. (2008).

19. Romero, Z. et al.  $\beta$ -globin gene transfer to human bone marrow for sickle cell disease. *J Clin Invest* 123, 3317–3330 (2013).

20. Pawliuk, R. et al. Correction of sickle cell disease in transgenic mouse models by gene therapy. 294, 2368–2371 (2001).
21. May, C., Rivella, S., Chadburn, A. & Sadelain, M. Successful treatment of murine beta-thalassemia intermedia by transfer of the human beta-globin gene. Blood 99, 1902–1908 (2002).
22. Negre, O. et al. Preclinical evaluation of efficacy and safety of an improved lentiviral vector for the treatment of beta-thalassemia and sickle cell disease. 15, 64–81 (2015).
23. Perumbeti, A. et al. A novel human gamma-globin gene vector for genetic correction of sickle cell anemia in a humanized sickle mouse model: critical determinants for successful correction. Blood 114, 1174–1185 (2009).
24. Gene Therapy for Sickle Cell Disease :A Lentiviral Vector Comparison Study. 29, 1153–1166 (2018).
25. Kumar, M., Keller, B., Makalou, N. & Sutton, R. E. Systematic Determination of the Packaging Limit of Lentiviral Vectors. 12, 1893–1905 (2001).
26. Canté-Barrett, K. et al. Lentiviral gene transfer into human and murine hematopoietic stem cells: size matters. 9, 312 (2016).
27. Peterson, K. R. & Stamatoyannopoulos, G. The 5' HS4 core element of the human  $\beta$ -globin locus control region is required for high-level globin gene expression in definitive but not in primitive erythropoiesis. 312, 17–26 (2001).
28. Lisowski, L. & Sadelain, M. Locus control region elements HS1 and HS4 enhance the therapeutic efficacy of globin gene transfer in  $\beta$ -thalassemic mice. 110, 4175–4178 (2007).

29. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *489*, 57–74 (2012).
30. Romero, Z. et al. The human ankyrin 1 promoter insulator sustains gene expression in a  $\beta$ -globin lentiviral vector in hematopoietic stem cells. *Molecular Therapy - Methods & Clinical Development* **2**, 15012 (2015).
31. Ryan, T. M., Ciavatta, D. J. & Townes, T. M. Knockout-Transgenic Mouse Model of Sickle Cell Disease. *Science* **278**, 873–876 (1997).
32. Levasseur, D. N., Ryan, T. M., Pawlik, K. M. & Townes, T. M. Correction of a mouse model of sickle cell disease: lentiviral/antisickling beta-globin gene transduction of unmobilized, purified hematopoietic stem cells. *Blood* **102**, 4312–4319 (2003).
33. Levasseur, D. N. & Townes, T. M. A recombinant human hemoglobin with anti-sickling properties greater than fetal hemoglobin. *J. Biol. Chem.* **279**, 27518–27524 (2004).
34. Grosveld, F., van Assendelft, G. B., Greaves, D. R. & Kollias, G. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* **51**, 975–985 (1987).
35. Ryan, T. M. et al. A single erythroid-specific DNase I super-hypersensitive site activates high levels of human beta-globin gene expression in transgenic mice. **3**, 314–323 (1989).
36. Effect of deletion of 5'HS3 or 5'HS2 of the human beta-globin locus control region on the developmental regulation of globin gene expression in beta-globin locus yeast artificial chromosome transgenic mice. *PNAS* **93**, 6605–6609 (1996).

37. Tanimoto, K. et al. Human  $\beta$ -globin locus control region HS5 contains CTCF-and developmental stage-dependent enhancer-blocking activity in erythroid cells. *23*, 8946–8952 (2003).
38. Fedosyuk, H. & Peterson, K. R. Deletion of the human beta-globin LCR 5'HS4 or 5'HS1 differentially affects beta-like globin gene expression in beta-YAC transgenic mice. *Blood Cells Mol. Dis.* *39*, 44–55 (2007).
39. Peterson, K. R. & Fedosyuk, H. LCR 5' hypersensitive site specificity for globin gene activation within the active chromatin hub. *40*, 11256–11269 (2012).
40. Forrester, W. C., Novak, U., Gelinias, R. & Groudine, M. Molecular analysis of the human beta-globin locus activation region. *PNAS* *86*, 5439–5443 (1989).
41. Collis, P., Antoniou, M. & Grosveld, F. Definition of the minimal requirements within the human beta-globin gene and the dominant control region for high level expression. *EMBO J* *9*, 233–240 (1990).
42. Fraser, P., Hurst, J., Collis, P. & Grosveld, F. DNaseI hypersensitive sites 1, 2 and 3 of the human  $\beta$ -globin dominant control region direct position-independent expression. *Nucleic Acids Res* *18*, 3503–3508 (1990).
43. Fraser, P., Antoniou, M., Grosveld, F. & Pruzina, S. Each hypersensitive site of the human beta-globin locus control region confers a different developmental pattern of expression on the globin genes. *Genes Dev.* *7*, 106–113 (1993).
44. May, C. et al. Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. *Nature* *406*, 82–86 (2000).

45. Pembrey, M. E., Wood, W. G., WEATHERALL, D. J. & Perrine, R. P. Fetal haemoglobin production and the sickle gene in the oases of Eastern Saudi Arabia. *Br. J. Haematol.* 40, 415–429 (1978).
46. Platt, O. S. et al. Mortality in sickle cell disease. Life expectancy and risk factors for early death. 330, 1639–1644 (1994).
47. Forget, B. G. Molecular basis of hereditary persistence of fetal hemoglobin. 850, 38–44 (1998).
48. Akinsheye, I. & Steinberg, M. H. Fetal hemoglobin in sickle cell anemia. *Blood* 118, 19–27 (2011).
49. Bao, E. L. et al. Heritability of fetal hemoglobin, white cell count, and other clinical traits from a sickle cell disease family cohort. *Am J Hematol* (2019).  
doi:10.1002/ajh.25421
50. Baglioni, C. & Weatherall, D. J. Abnormal human hemoglobins. IX. Chemistry of hemoglobin J-Baltimore. *Biochim Biophys Acta* 78, 637–643 (1963).
51. Dull, T. et al. A third-generation lentivirus vector with a conditional packaging system. 72, 8463–8471 (1998).
52. Hindson, B. J. et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. 83, 8604–8610 (2011).
53. Hope, T. in *Lentiviral Vectors* 261, 179–189 (Springer, Berlin, Heidelberg, 2002).
54. Antoniou, M., Geraghty, F., Hurst, J. & Grosveld, F. Efficient 3'-end formation of human beta-globin mRNA in vivo requires sequences within the last intron but occurs independently of the splicing reaction. *Nucleic Acids Res* 26, 721–729 (1998).

55. Behringer, R. R., Hammer, R. E., Brinster, R. L., Palmiter, R. D. & Townes, T. M. Two 3' sequences direct adult erythroid-specific expression of human beta-globin genes in transgenic mice. *PNAS* 84, 7056–7060 (1987).
56. Antoniou, M., deBoer, E., Habets, G. & Grosveld, F. The human beta-globin gene contains multiple regulatory regions: identification of one promoter and two downstream enhancers. *EMBO J* 7, 377–384 (1988).
57. Gelinis, R., Frazier, A. & Harris, E. A normal level of beta-globin expression in erythroid cells after retroviral cells transfer. *Bone Marrow Transplant* 9 Suppl 1, 154–157 (1992).
58. Leboulch, P. et al. Mutagenesis of retroviral vectors transducing human beta-globin gene and beta-globin locus control region derivatives results in stable transmission of an active transcriptional structure. *EMBO J* 13, 3065–3076 (1994).
59. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res* 33, W526–W531 (2005).

**Chapter 3: Next Generation  $\beta$ -globin-Expressing Lentiviral Vectors Created Using  
a Massively Parallel Reporter Assay – Applying Rapid Dissection of a Locus  
Control Region to Accelerate Gene Therapy Vector Development**

## **Abstract**

Hematopoietic stem cell gene therapy is a promising approach in treating disorders of the hematopoietic system. Identifying combinations of cis-regulatory elements that do not impede packaging or transduction efficiency when included in lentiviral vectors has proven challenging. Candidate enhancers or promoters must be tested against a litany of performance criteria until high-performing combinations can be found. Here we deploy LV-MPRA (Lentiviral Vector-based, Massively Parallel Reporter Assay), an approach that simultaneously analyzes thousands of synthetic DNA fragments in parallel to identify sequence intrinsic and lineage-specific enhancer function at near base pair resolution. We demonstrate the power of LV-MPRA in elucidating the boundaries of previously unknown intrinsic enhancer sequences of the globin Locus Control Region. Our approach facilitated the rapid assembly of a novel therapeutic vector harboring a strong lineage-specific recombinant control element capable of correcting a mouse model of Sickle Cell Disease. LV-MPRA can be used to map any genomic locus for enhancer activity and facilitates the rapid development of therapeutic vectors for treating disorders of the hematopoietic system or other specific tissues and cell types.

## **Introduction**

Hematopoietic stem cell gene therapy is a promising approach for treating many monogenic disorders of the hematopoietic system, having demonstrated remarkable success in a number of phase I/II clinical trials (Aiuti et al., 2013; Biffi et al., 2013; Boelens et al., 2013; Cooper et al., 2017; Hacein-Bey-Abina et al., 2015; 2010; Ribeil et

al., 2017; Sessa et al., 2016; Thompson et al., 2018; Walters, 2015). Therapeutic lentiviral vectors typically contain human cis-regulatory elements that provide high-level lineage-specific expression of a therapeutic transgene. Identification of appropriate cis-regulatory elements that do not impede packaging or transduction efficiency when included in lentiviral vectors (LVs) has been a major challenge in developing novel gene therapy vectors (Delzor et al., 2012; Lisowski and Sadelain, 2007; May et al., 2000; Molette et al., 2001; Sadelain et al., 1995; Weber et al., 2018; Zhan et al., 2016).

The elements used in LVs to confer tightly regulated expression of therapeutic transgenes are often identified by contrasting mouse knock-out studies to histone marking and accessibility data (The ENCODE Project Consortium, 2012) to determine putative boundaries of cis-regulatory elements. While aforementioned technologies have enabled the identification of candidate cis-regulatory elements for incorporation into LVs, these technologies fail to provide insight on the exact base pair boundaries of bona-fide “sequence intrinsic” enhancers (i.e. the actual sequences providing enhancer activity). Moreover, when chromatin marking and accessibility data are used to identify putative regulatory elements, detection of appropriate cell-type/state specific control elements can be confounded by extrinsic factors such as chromatin state, cell-cycle, cell-state, cell-type, antibody sensitivity/specificity, protein evanescence, etc. The proposed control sequences must still be functionally tested for their ability to enhance transgene expression, which is a low throughput process when attempting to identify the best combination of elements to include in LVs.

To overcome the limitations of current methods used to detect cell-specific enhancer elements, we developed a Massively Parallel Reporter Assay (MPRA) to

identify human genomic sequences capable of driving lineage-specific reporter gene expression in a lentiviral vector (LV) system. Similar assays have been used to map promoter activity (van Arensbergen et al., 2017), measure enhancer activity (Inoue et al., 2017; 2018; Kwasnieski et al., 2014; Maricque et al., 2017), and have provided novel insights into the complex relationships between nucleotide composition and regulatory activity (Grossman et al., 2017; Kreimer et al., 2019; Smith et al., 2013; White et al., 2013; Yanez-Cuna et al., 2014).

Here, we show that LV-MPRA was able to map the functional cis-regulatory elements possessing enhancer activity— at near base pair resolution – of the nearly 20kb or so  $\beta$ -Globin LCR, a region known to contain potent lineage-specific enhancers that drive the temporal and developmentally specific expression of the  $\beta$ -globin gene locus (Philipsen and Hardison, 2018). The sequence intrinsic enhancer map generated by LV-MPRA guided the construction of an array of novel recombinant regulatory elements that conferred a range of transgene expression in human erythrocytes that correlated with the length of concatenated enhancer combinations.

Finally, we demonstrated that novel therapeutic  $\beta$ -globin LV constructs generated using LV-MPRA data, selected for optimal size and expression, faithfully corrected the disease phenotype in a mouse model of Sickle Cell Disease (SCD), demonstrating the application of this LV-MPRA toward developing clinically relevant therapeutic vectors.

## RESULTS

### LV-MPRA Identifies Sequence Intrinsic Enhancers of B-Globin Locus Control

#### Region

In order to identify enhancer regions, we required a reporter vector that contained only the minimal human  $\beta$ -Globin promoter as the major cis-regulatory element driving expression of the  $\beta$ AS3-P2A-mCIT transgene {Levasseur:2004dg, Donnelly:2001cz}. For proof of principle, we incorporated a 103 bp sequence derived from the HS2 element of the  $\beta$ -globin LCR into the multiple cloning site LV-reporter (Supplementary Figure 1A), this provided increased expression of the reporter gene in Human Umbilical Cord Blood-Derived Erythroid Progenitor Clone 2 cells (HUDEP-2) (Kurita et al., 2013) when compared to controls (Supplementary Figure 1B and 1C). This finding suggested that ~100bp regions of the LCR could enhance expression of the minimal human  $\beta$ -Globin promoter. Thus, LV-Reporter became the basis of what was used to analyze thousands of overlapping ~100bp LCR sequences in parallel for enhancer activity.

The LV-MPRA experimental strategy consisted of three steps: 1) *in silico* library design and construction, 2) library packaging, cell line transduction and culture, 3) barcode acquisition, sequencing, and data analysis.

First, a library of overlapping 103 bp LCR sequences were designed by *in silico* tiling. Overlapping sequences were generated with the start of subsequent sequences beginning 4 bp after the start of a preceding sequence. The collection of sequences were duplicated three times by associating three unique 13 bp barcodes with each sequence. The resultant sequence collection was duplicated a second time by replacing all sequences with reverse complement sequences and assigning new barcodes. In

total a collection of ~25,000 unique LCR/barcode sequence combinations were generated with a maximum of 150x coverage per base pair (Figure 1A).

The library of LCR sequence and barcode pairs were synthesized via DNA microarray as 170 bp oligonucleotides. A schematic of a single 170mer is shown in Figure 1B. A cloning strategy was deployed that allowed for placement, *en masse*, of LCR sequences upstream of the minimal  $\beta$ -globin promoter, and placement of corresponding barcodes downstream of the  $\beta$ AS3-P2A-mCIT sequences and upstream of the 3'UTR within the vector's transcriptional cassette. Thus, the strength of the query sequence could be quantified by the abundance of barcodes expressed in mRNA.

Second, the plasmid library was packaged into viral particles and used to transduce the HUDEP-2 cell line. To verify that packaging and cell line transduction did not negatively influence library complexity, we quantified barcode abundance in the starting plasmid pool and genomic DNA of transduced HUDEP-2 cells. The correlation between barcodes in the plasmid pool and genomic DNA of transduced HUDEP-2 cells was strong ( $r=0.854$ ), demonstrating that the diversity of barcodes in the starting plasmid pool was efficiently transferred to the integrated proviral barcode pool of transduced HUDEP-2 cells (Supplementary Figure 2A; each dot in the dot plot represents the  $\log_{10}$  value of an individual barcode found in the plasmid barcode pool, integrated proviral barcode pool, or both).

Lastly, barcodes were obtained from mRNA of transduced HUDEP-2 cells and sequenced after cells were differentiated down the erythroid lineage for four days. The correlation between mRNA and gDNA barcodes was low ( $r=0.304$ ), demonstrating that only a subset of "Query" sequences possessed enhancer activity (Supplementary

Figure 2B). To generate a map of enhancer activity across the LCR, barcode reads were normalized by sequencing depth and enhancer activity calculated by dividing RNA barcode counts by plasmid DNA barcode counts. Plasmid barcode counts were chosen as the normalization factor as we were only interested in identifying enhancer sequences that were successfully packaged to be present to enhance expression in the HUDEP-2 cells. Normalized barcode reads were generated for each starting position and plotted to their corresponding positions across a map of the LCR.

The resulting map of enhancer activity across the LCR is provided in Figure 2A. Highlighted regions represent boundaries of LCR regulatory elements characterized in transgenic mice as erythroid cell-specific DNase I hypersensitive sites that were needed for high level position-independent expression (Collis et al., 1990; Fedosyuk and Peterson, 2007; Fraser et al., 1993; 1990; Lisowski and Sadelain, 2007; Peterson et al., 1996; Tanimoto et al., 2003) . As shown, peaks within the LCR that represent sequences with strong intrinsic enhancer activity fall within the boundaries of “classically” defined LCR control elements. The same map of sequence intrinsic enhancer activity aligned to ENCODE track sets denoting regions of open chromatin and other markings (markings that when present in specific combinations suggest the presence of enhancers (King et al., 2016) is shown (Supplementary Figure 2D).

To determine if the presence of specific transcription factor binding sites (TFBS) were predictive of intrinsic enhancer strength, the presence of key putative erythroid specific TFBS (GATA1, KLF1, and TAL1) were identified and aligned to the map of normalized barcode counts (Figure 2B). The presence of putative erythroid specific TFBS were seen to align closely with intrinsic enhancer peaks above the 95th percentile

of highest expression. These findings are consistent with previous reports that highlight the role of GATA1, KLF1, and TAL1 in modulating expression of erythroid-specific genes (Kang et al., 2015; Tallack et al., 2012).

Densities of putative TFBS were also found to be predictive of enhancer strength. Putative TFBS were identified along the LCR using FIMO (Grant et al., 2011), an algorithm that scans a given DNA region for individual TFBS consensus sequence matches. Figure 2C provides a density map of FIMO TFBS matches plotted by location across the LCR. Locations with quantities of putative TFBS greater than 40 were seen to align closely with sequence intrinsic enhancer peaks that were above the 95th percentile of highest expression.

### **LV-MPRA Guided Therapeutic Vector Design and Characterization**

A series of  $\beta$ -globin expression constructs were designed by imposing thresholds to cluster sequences that were within the 80<sup>th</sup>, 90<sup>th</sup>, 95<sup>th</sup>, 97.5<sup>th</sup>, or 98.75<sup>th</sup> percentiles of highest enhancer activity (Figure 3A). Concatenated enhancer combinations were each cloned in sense orientation into the identical  $\beta$ -globin expression vector backbone and compared head-to-head against  $\beta$ AS3LV, which was deemed to be the strongest expressing globin-vector by way of its full-length HS 2, 3, and 4 elements (totaling ~3.6kb in sequence length).

The LV-MPRA guided constructs, termed 80, 90, 95, 97.5, and 98.75 (based on the percentile cutoff thresholds used to identify sequences for concatenation within a given percentile of expression), were packaged in parallel against  $\beta$ AS3LV using HEK 293T cells and titered head-to-head on HT-29 cells. An inverse correlation was

observed when titer was plotted by function of enhancer length, demonstrating that overall enhancer length strongly influenced infectious particle production (Figure 3B).

To measure the infectivity of the different vectors, human CD34<sup>+</sup> hematopoietic stem and progenitor cells (HSPCs) were transduced at a fixed multiplicity of infection [MOI] of 10 ( $1.0 \times 10^7$  TU/mL (based on titers measured on HT-29 cells). Transduced CD34<sup>+</sup> cells were cultured for 14 days followed by extraction of gDNA for measurement of vector copy number (VCN). An inverse relationship was observed between enhancer length and VCN, which reflects relative infectivity of the different vectors (Figure 3C and Table 1).

Transduction efficiency was then assessed across a range of MOIs (1, 3.3, 6.6, or 10). Linear regression analysis revealed a positive correlation between infectivity and vector dose while demonstrating an inverse correlation between infectivity and enhancer length (Figure 3D).

We then sought to examine the relationship between enhancer length and expression per vector genome. Thus, HSPCs were transduced at various MOIs and cultured under erythroid differentiation conditions for 14 days. Expression by each vector was measured as percentages of  $\beta^{\text{AS3}}$ -globin transcripts to total  $\beta$ -globin-like transcripts, normalized to VCN, and plotted by function of enhancer length. Normalized expression levels were seen to correlate strongly with enhancer length (Figure 3E).

Comparing % $\beta^{\text{AS3}}$ -globin transcript levels and VCN, each vector had a characteristic linear relationship (Figure 3F). The larger vectors achieved higher expression per VCN, but the smaller vectors were able to reach higher VCN and compensate for lower expression per VCN to yield total amounts of expression that

were similar. For example,  $\beta$ AS3LV was seen to provide slightly higher levels of expression than did similarly sized 80 vector (as shown by the increased slope of the fitted line), but this difference failed to reach significance. All LV-MPRA vectors displayed erythroid-specific expression patterns, as expression was non-detectable in transduced HSPCs cultured under myeloid differentiation conditions (Table 1).

We then sought to determine optimal enhancer orientation and arrangement. Sequences within the 95<sup>th</sup> percentile of highest expression were arranged in sense orientation with those associated with HS1 placed closest to the promoter (95-Sense). Enhancers were also arranged in an anti-sense orientation, with those elements associated with HS1 placed furthest from the promoter (95-Antisense). Enhancers were also arranged in anti-sense orientation with those associated with HS1 now placed closest to the promoter (95-Alt Antisense). A schematic outlining enhancer orientation and arrangement is provided in Supplemental Figure 3A. The collection of enhancer configurations were cloned into the identical  $\beta$ -globin expression vector backbone and packaged and titered in parallel. No significant differences in titer were observed (Supplemental Figure 3B). Raw vector supernatants were then used to transduce HUDEP-2 cells at MOI 1 and cells cultured and differentiated for 14 days. No differences in infectivity (Supplemental Figure 3C), expression (Supplemental Figure 3D), or normalized expression levels per VCN were observed, demonstrating that element directionality or element proximity to the promoter did not diminish vector production, transduction or expression (Supplemental Figure 3E).

Lastly, we sought to confirm that sequences within the 95<sup>th</sup> percentile of lowest expression were incapable of enhancing expression above basal levels offered by the

minimal promoter. Sequences within the 95<sup>th</sup> percentile of lowest expression were identified, concatenated and cloned into a  $\beta$ -globin expression vector backbone to generate 95-Negative. This new construct was then compared against a  $\beta$ -globin expression vector containing only a minimal  $\beta$ -globin promoter as the major regulatory element driving expression (Pro.Only) and the previously defined (above) 95-sense construct. Human CD34+ HSPCs were transduced at MOI 10 and culture under erythroid conditions for 14 days. The previously defined 95-sense construct offered significantly higher levels of normalized expression when compared head-to-head against other constructs. As expected, no significant differences in normalized expression between 95-Negative and Pro.Only were observed, demonstrating that sequences within the 95<sup>th</sup> percentile of lowest expression were incapable of enhancing expression while sequence within the 95<sup>th</sup> percentile of highest expression were capable of enhancing expression (Supplementary Figure 4C).

Taken together, these data demonstrate that cis-regulatory elements identified using LV-MPRA fulfill the major tenet that defines enhancers, which is the ability to increase transcription from a promoter independently of direction or juxtaposition (Pennacchio et al., 2013; Pott and Lieb, 2015). These data also demonstrate that LV-MPRA can be used to concatenate regulatory elements of various lengths that (in this case) offer predictable patterns of performance across multiple categories (additional data provided in Supplementary Figure 5).

## ***In Vivo* Characterization of LV-MPRA Based Therapeutic Vectors in SCD Mouse**

### **Model**

The best performing LV-MPRA constructs, 95 and 97.5, which were intermediate for length, titer and expression, were then compared to Lenti/ $\beta$ AS3-FB in the “Townes” mouse model of SCD to evaluate their ability to induce hematologic correction. Lineage-depleted bone marrow cells were obtained from homozygous  $\beta^S/\beta^S$  donor mice and pre-stimulated for one-day. Cells were transduced at equal MOI by the different vectors and after one day delivered by retro-orbital injection into lethally irradiated GFP-transgenic mouse recipients (B6-GFP; Jackson Laboratory). Peripheral blood (PB) samples, acquired 4- and 16-weeks post-transplantation, were assessed for engraftment by flow cytometry for GFP-negative donor cells, gene marking in circulating cells (VCN by ddPCR), and blood hemoglobin concentration and composition by high-performance liquid chromatography (HPLC).

Mice with bone marrow (BM) donor engraftment <97% at week 16 were excluded from analyses as  $\geq 4\%$  residual WT recipient RBCs could mask adverse pathophysiology induced by  $\beta^S/\beta^S$  donor cells (Levasseur et al., 2003). Week 16 engraftment efficiency in PB or BM (by FACS or HPLC) was not different among experimental arms (Supplementary Figure 6A-C). Average gene transfer efficiency seen in circulating PB (Figure 4A) and BM cells (Supplementary Figure 6D) differed significantly among experimental arms, and demonstrated that constructs with decreased total enhancer lengths offered superior transduction efficiency.

Quantification of hemoglobin (HB)  $\beta^{AS3}$  tetramers in peripheral blood lysates was accomplished using HPLC. To compare differences in normalized expression between

each experimental arm, %Hb $\beta^{AS3}$ /total hemoglobin tetramers were normalized to PB VCN for each mouse and plotted. Constructs with larger enhancers offered superior expression per vector genome with overall trends reflecting those seen in in vitro experiments, where constructs with larger enhancers offered superior expression per vector genome (Figure 4C).

Surprisingly, average total levels of Hb $\beta^{AS3}$ /total hemoglobin tetramers were not different among experimental arms demonstrating that lower expression levels from smaller vectors were compensated for by gains in transduction efficiency (Figure 4B). Expression findings were also confirmed by measuring % $\beta^{AS3}$ -globin mRNA to  $\beta$ -globin mRNA in BM obtained 16 weeks post transplantation and then normalizing those values to BM VCN (Supplemental Figure 6C).

Finally, hematologic indices were measured using PB obtained 4- and 16- weeks post transplantation. When compared to mice transplanted with mock-transduced BM cells, hemoglobin levels, red blood cell counts, and hematocrits were significantly higher in mice that were transplanted with BM transduced with vectors designed using LV-MPRA.

At week 16, Hb levels of mice that received mock-transduced cells were 7.9 g/dL on average while the Hb levels of mice that received BM transduced with Lenti/ $\beta^{AS3}$ -FB, 95 or 97.5 were 11.0, 11.2 and 11.6 g/dL on average, respectively (Figure 4D). RBC counts were also significantly higher for recipients of Lenti/ $\beta^{AS3}$ -FB, 95 or 97.5 transduced BM cells ( $8.5, 9.0 \times 10^6$  and  $8.5 \times 10^6$  cells/ $\mu$ l on average, respectively) compared to recipients of mock-transduced BM cells ( $6.2 \times 10^6$  cells/ $\mu$ l on average). Similar improvements were seen for hematocrits, where mice that received mock-

transduced cells had hematocrits of 25.2 on average, while mice that received Lenti/ $\beta$ AS3-FB, 95 or 97.5 transduced BM had hematocrits of 30.9, 32.2 and 30.6 on average, respectively. The *in vivo* data demonstrate that when LV-MPRA was deployed towards therapeutic vector design, identification of correct combinations of enhancers capable of providing sufficient levels of transgene expression was achieved.

## **Discussion**

Control over lentiviral vector transgene expression is typically achieved by including cis-regulatory elements that enhance transcription of a LV's internal promoter in a cell type-specific pattern (Morgan et al., 2017). While genome-scale cell-line based genetic and epigenetic studies have generated an impressive collection of candidate enhancers, these studies utilize indirect measurements (histone modification, chromatin accessibility, and bound transcriptional co-activators) to predict their locations (Zhang et al., 2016). Moreover, the limitations of these studies fail to provide exact boundaries of sequence intrinsic enhancers (i.e. the actual sequences that provide enhancer activity). To-date, the vast majority of putative cis-regulatory elements have yet to be characterized in lentiviral vector systems.

Therapeutic expression vectors are traditionally designed by testing individual candidate enhancers for their ability to potently drive cell-type specific transgene expression, in addition to determining how incorporation influences titer and infectivity. While criteria can be imposed to reduce the number of candidate enhancer elements for evaluation, only a handful of elements can be tested at a time as current throughput limitations restrict the number of LVs that can be evaluated in parallel. Confounding the

issue of low throughput, specific combinations of elements must also be uncovered to achieve temporal and spatial control over transgene output. Thus, the challenges to developing expression vectors produced at high-titer, while offering both robust infectivity and appropriate levels of cell-type specific protein expression, are elevated.

Multiple groups have deployed high-throughput approaches toward dissecting the functional activities of cis-regulatory elements in various context (Grossman et al., 2017; Inoue et al., 2017; 2018; Kreimer et al., 2019; Kwasnieski et al., 2014; Maricque et al., 2017; Smith et al., 2013; van Arensbergen et al., 2017; White et al., 2013; Yanez-Cuna et al., 2014); however, no group has applied their findings towards therapeutic vector design. While each group developed an assay that possessed aspects of an ideal functional assay for therapeutic vector development, each approach had its own unique advantages and limitations. The ideal functional assay is one that can unbiasedly screen thousands of DNA sequences derived from large genomic regions for intrinsic enhancer activity and do so in the context of a therapeutic vector integrated into a relevant cell type.

To this end, we developed LV-MPRA to dissect at high-resolution the sequence intrinsic enhancers of the  $\beta$ -globin locus control region in an erythroid progenitor cell line. By using DNA oligonucleotide synthesis to create the starting material required to create thousands of barcoded LVs each harboring a known LCR fragment, we were able to transcend above the traditional limitations of therapeutic vector development by using barcode sequencing to digitally measure enhancer activity across thousands of sequences. The LV-MPRA allowed us to generate a continuous and quantitative map on which to dissect the ability of putative cis-regulatory sequences to enhance

expression while providing exact boundaries of sequence-intrinsic enhancers within the  $\beta$ -globin LCR.

When the percentile cutoff of 80% was imposed to identify sequences within the 80th percentile of highest expression, we were able to develop a novel regulatory sequence that when incorporated into a therapeutic expression vector backbone produced a construct, termed 80, with no significant differences in titer, infectivity, or expression when compared to Lenti/ $\beta$ AS3-FB. Given that boundaries defining the enhancer elements of Lenti/ $\beta$ AS3-FB have been undergoing continuous refinement for the past 20-years (Weber et al., 2018), we found it noteworthy that our approach was able to create a comparable regulatory sequence of novel composition in a matter of weeks.

Since the 80 vector was so analogous to Lenti/ $\beta$ AS3-FB (in relation to robust expression but diminished titer and gene transfer), we decided to test the smaller 95 and 97.5 in the “Townes” mouse model of SCD. Although 95 and 97.5 were found to possess 50% and 30% of the expression activity of Lenti/ $\beta$ AS3-FB per vector genome, respectively, these losses in expression were overcome by gains in infectivity resulting in production of comparable total amounts of  $\beta$ AS3-mRNA and protein in circulating murine RBCs, thus reversing the disease phenotype. While an increase in VCN to compensate for expression is typically undesirable, gene transfer is typically <1 copy per cell in the case of traditional  $\beta$ -globin-expressing lentiviral vectors (without the use of transduction enhancers). Thus, an increase to a VCN of 2-4 is not excessive, as seen with the 95 vector. More importantly, 95 and 97.5 were found to be produced at

significantly higher titer when compared to Lenti/ $\beta$ AS3-FB, which may hold implications for improving access to gene therapy by expanding the vector supply.

This proof of principle study demonstrates that in mere weeks, lineage-specific therapeutic expression vectors were generated that rivalled a vector that has undergone years of refinement. Designing lentiviral vectors using “sequence intrinsic” enhancers of the LCR was done not just because the locus is such a hot gene therapy target, but also because fine-mapping represents a high bar to clear since the locus has been studied in such detail for decades. The real power of this LV-MPRA approach should come in to play when far less studied loci are targeted for gene therapy vector design. This framework can be deployed to develop any therapeutic expression vector for any specific cell type and advances the prospect of realizing the promise of gene therapy.

## **Methods**

### **LV-MPRA oligo design and preparation**

Pools of synthetic 170-mer DNA sequences were ordered from CustomArray Inc. (Bothell, WA). Each oligonucleotide was designed according to the following scheme: 5' primer amplification sequence (ATGTTTTTCTAGGTCTCGAG)/ 103 bp LCR “query” sequence/ *BmtI* site/ 4 bp spacer/ *Sall* site/Barcode/ 3' primer amplification sequence (ctttgtccctaagtccaac).

Each oligonucleotide possessed a 103 bp “query” sequences derived from a larger ~16KB LCR sequence, with each subsequent oligonucleotide possessing a similar “query” sequence offset by 4 bp until complete coverage of the larger sequence was achieved. A total of  $\sim 4.2 \times 10^3$  unique oligos were needed to achieve 1x coverage of

the larger 16kb LCR sequence. Each “query” sequence was assigned three unique 13 bp barcodes tripling the diversity of sequences to  $\sim 1.2 \times 10^4$ . Antisense versions of the query sequences were also included, doubling the total number of unique “query”/barcode pairs to  $\sim 2.5 \times 10^4$ .

Primer “Custom Array Rev” (Supplemental Table 2) was used to convert full-length DNA microarray synthesized oligonucleotides into dsDNAs by primer extension. Twelve reactions using MyTaqRed (Bioline Meridian Life Science, Inc., Memphis TN) were established in parallel with each reaction containing 200 ng of template and 4  $\mu$ l of REV primer. A single cycle of PCR amplification was performed at annealing temperature of 45°C. Reaction products were purified using PureLink™ PCR Purification Kit (Thermo Fisher Scientific, Waltham MA) to deplete residual primers and ssDNA products and pooled.

Primers “CustomArray FWD” and “CustomArray REV” were then used to amplify 170 bp dsDNA fragments. Twelve PCR reactions were performed using MyTaqRed (Bioline Meridian Life Science, Inc., Memphis TN) with 200 ng of dsDNA template and amplified for 6-cycle with annealing temperature of 45°C. Reaction products were purified by column and pooled. The purified reaction products were then used in library construction.

### **LV-MPRA library construction**

We created a streamlined version of a globin expression vector for use in LV-MPRA library construction. The modified construct was made by replacing the reverse oriented human genomic elements (full-length LCR HS 2 and 3 enhancer elements,

minimal  $\beta$ -globin gene promoter,  $\beta$ AS3 transgene with introns, exons, and 3'UTR) of Globe1-AS3-FB with a minimal  $\beta$ -globin gene 3'UTR sequence and an *EcoRV* site. The plasmid backbone was linearized using *EcoRV* enzyme in Buffer 3.1 (New England Biolabs, Ipswich, MA) and 5' DNA ends dephosphorylated using recombinant Shrimp Alkaline Phosphatase (New England Biolabs, Ipswich, MA). Reaction products were pooled and run on 1% agarose gel and DNA of correct length purified gel extracted and purified using PureLink™ Quick Gel Extraction Kit (Invitrogen, Carlsbad, CA, United States). Pools of 170 bp dsDNA fragments were cloned into linearized plasmid backbones using New England Biolabs' NEBuilder HiFi DNA Assembly (New England Biolabs, Ipswich, MA) mix to create the "Pre-complete library" (the 5' and 3' primer amplification sequences of the 170 bp dsDNA fragments had homology with the DNA ends of the linearized plasmid backbones). A total of six 20  $\mu$ l NEBuilder reactions were established, each containing 400 ng of purified plasmid backbone, 50ng of purified 170 bp dsDNA fragments. Reactions were incubated at 50°C for 60 minutes. The NEBuilder reactions were pooled, purified using 2.5X Agencourt AMPure XP Beads Kit (Beckman Coulter, Brea, CA, USA), and eluted into a final volume of 24  $\mu$ l DNase free water. A total of six transformation reactions were performed with Stable Competent E.coli (New England Biolabs, Ipswich, MA) and 4  $\mu$ l of purified plasmid product. Each transformation was recovered in 500  $\mu$ l of SOC for 30 minutes, pooled, and expanded in 500 ml of LB overnight. Large scale plasmid DNA isolation was then performed using PureLink™ HiPure Plasmid Maxiprep Kit (Invitrogen, Carlsbad, CA, United States) to create "Pre-Complete Library"

The “Pre-Complete Library” was then linearized with *Sall* and *BmtI* enzymes in Buffer 3.1 (New England Biolabs, Ipswich, MA). A total of 12 digestion reactions were established each containing 2ug of plasmid. Digest were incubated over night at 37°C with recombinant Shrimp Alkaline Phosphatase. Reaction products were pooled and run on 1% agarose gel and DNA of correct length gel extracted and purified.

The  $\beta$ AS3-P2A-MCIT expression cassette was liberated from its pCR-BluntII-TOPO vector backbone using *Sall* and *BmtI* restriction enzymes in Buffer 3.1 (New England Biolabs, Ipswich, MA). A total of 12 digestion reactions were established each containing 2  $\mu$ g of plasmid. Digest were incubated over night at 37°C. Reaction products were pooled and run on 1% agarose gel and DNA of correct length gel extracted and purified.

The expression cassette was then ligated into the linearized “Pre-Complete Library”. A total of six ligation reactions were established each containing 200 ng of linearized plasmid and 430 ng of insert. The ligation reactions were pooled, purified using 2.5X Agencourt AMPure XP Beads Kit (Beckman Coulter, Brea, CA, USA) and eluted into a final volume of 24  $\mu$ l. A total of six transformation reactions were performed with Stable Competent E.coli (New England Biolabs, Ipswich, MA). and 4  $\mu$ l of purified plasmid product. Each transformation was recovered in 500  $\mu$ l of SOC for 30 minutes, pooled, and expanded in 500ml of LB overnight. Large scale plasmid DNA isolation was then performed using PureLink™ HiPure Plasmid Maxiprep Kit (Invitrogen, Carlsbad, CA, United States) to create “Complete Library”

## **Vector production and titration**

Transient transfection of 293T cells using the third-generation packaging system (Dull et al., 1998) provided packaged virus particles. Viral supernatants were then directly used for titer determination or concentrated by tangential flow filtration, as described by Cooper et al (Cooper et al., 2011). Briefly, the HT-29 human colorectal carcinoma cell line was transduced with different dilutions of both raw and concentrated vectors. To calculate titers, cells were harvested and VCNs were determined by ddPCR approximately 60 hours post transduction.

## **HUDEP-2 cell culture and transduction**

HUDEP-2 cells (provided by Dr. Nakamura, RIKEN BioResource Center, Tsukuba, Ibaraki, Japan) were maintained in Iscove's modified Dulbecco's medium (IMDM; Gibco, Grand Island, NY, USA) supplemented with 1  $\mu$ M dexamethasone (Sigma–Aldrich, St Louis, MO), 1  $\mu$ g/ml doxycycline (Sigma–Aldrich, St Louis, MO), 50 ng/ml human SCF, 3 units/ml EPO (all cytokines were acquired from PeproTech, Rocky Hill, NJ), and 1x glutamine, penicillin, and streptomycin (Gemini Bio-Products, Sacramento, CA). A total of 2.5e6 HUDEP-2 cells were transduced with 5ml of raw virus in a total volume of 10ml in a T75 Flask. Approximately 24 hours later, cells were pelleted, medium changed and cells expanded for 7 days. Approximately  $1.0 \times 10^7$  Transduced HUDEP-2 cells were then harvested for gDNA isolation and remaining cells plated on an MS5 stromal cell layer in IMDM supplemented 1x glutamine, penicillin, and streptomycin, holo-human transferrin (330  $\mu$ g/ml, Sigma–Aldrich, St Louis, MO), heparin (2 IU/ml, Sigma–Aldrich, St Louis, MO), recombinant human insulin (10  $\mu$ g/ml, Sigma–

Aldrich, St Louis, MO), 3 units/ml EPO and 5% inactivated human plasma (Grifols USA, Los Angeles, CA). Cells were co-cultured for 4 days, after which  $\sim 1.0 \times 10^7$  cells were harvested for RNA extraction.

### **Barcode generation**

Genomic DNA (gDNA) was extracted from HUDEP-2 cells 8 days after transduction using PureLink™ Genomic DNA Mini Kit (Thermo Fisher Scientific, Carlsbad, CA, USA). Total RNA was extracted from HUDEP-2 cells 4 days after the start of differentiation using the RNeasy Mini Kit (Qiagen, Inc., Valencia, CA). Total RNA was separated into eight 10  $\mu$ l aliquots for storage. Reverse transcription of total RNA was carried out in parallel across six reactions each using 10  $\mu$ l of total RNA, 1  $\mu$ l of “1<sup>st</sup> Strand Primer”, 2  $\mu$ l of 0.1mM DDT, 4  $\mu$ l of First Strand Synthesis Buffer, 1  $\mu$ l of 10mM dNTP’s, 1  $\mu$ l of RNase out, and 1  $\mu$ l of 200U/ $\mu$ l of M-MLV reverse transcriptase (all from Invitrogen, Carlsbad, CA). The 1<sup>st</sup> strand primer was designed to enrich for only cDNAs that contained barcodes and thus, possessed homology to mCIT to allow for creation of cDNAs comprised of sequence spanning mCIT and the 3’ UTR of the barcoded  $\beta$ AS3-P2A-MCIT mRNAs. To amplify barcodes from both gDNA, RNA and “Complete Library” plasmid, primers 1<sup>st</sup> strand primer (nested)” and 2<sup>nd</sup> strand primer were used as they flank the barcoded region. Following PCR amplification, both the gDNA and cDNA barcodes were purified and submitted to the UCLA Technology Center for Genomics & Bioinformatics Sequencing Core, Department of Pathology & Laboratory Medicine for library construction and PEx150 Illumina sequencing. Reads that perfectly matched the first 14 nucleotides of the amplicon were included in subsequent analysis. We

generated 124 million reads from the “complete library” plasmid, and 117 million reads from the cDNA, and 49 million reads from the gDNA. Plasmid barcode reads were highly correlated with gDNA reads ( $r= 0.854$ ), allowing the plasmid barcode reads to be used for normalization.

### **Barcode quantification and sequencing**

The constant sequences before (ACTTAGGGGAACAAAG) and after (GTCGACATGCTAGC) the barcode were used to locate barcodes from sequencing data and the frequencies of barcodes determined in plasmid, gDNA and cDNA. To quantify the enhancing ability of each LCR fragment, the read counts were first normalized by sequencing depth, cDNA barcode counts were divided by plasmid barcode counts after the sums of matched barcode triplets were calculated (each LCR fragment was associated three unique barcodes). The  $\log(x+1)$  values were calculated and used to represent enhancing ability.

The LCR sequences tested were 103 bp in length with each neighboring sequence beginning 4 bp away from the next. Thus, a single nucleotide should be covered by 150 different sliding fragments. Statistical boot-strapping was used to calculate the mean of the non-zero counts in the 150 covering fragments to represent the enhancing score. The enhancing score was then used to make the map of enhancer activity across the LCR in Figure 2A.

To quantify putative transcription factor (TF) binding sites present within the LCR in the enhancer, FIMO 5.0.1 (Grant et al., 2011) was used to infer locations of TF binding sites across query sequence. A total of 771 motifs for human TFs from

HOCOMOCOv11 were input to FIMO for searching, and the p-value threshold set to 0.0001. The number of motifs binding to each nucleotide were calculated and plotted in Figure 2C.

### **BM CD34+ cell culture and transduction**

All bone marrow aspirates were obtained from voluntary healthy donors supplied by AllCells (Alameda, CA). BM mononuclear cells were isolated by Ficoll-Hypaque density gradient centrifugation. CD34+ HSPCs were enriched using CD34+ MicroBead Kit (Miltenyi Biotec, Bergisch Gladbach, Germany). Enriched CD34+ HSPCs were cryopreserved in Fetal Bovine Serum supplemented with 10% dimethyl sulfoxide (Sigma–Aldrich, St Louis, MO) in liquid nitrogen. Cells were thawed and plated on non-tissue culture-treated six-well plates pre-coated with RetroNectin (20 µg/ml, Takara Shuzo Co., Otsu Japan) at  $1 \times 10^6$  cells/mL. Cells were pre-stimulated for 16-24 hours in X-Vivo 15 medium (Lonza, Basel, Switzerland) supplemented with 1x glutamine, penicillin, and streptomycin (Gemini Bio-Products, Sacramento, CA), human stem cell factor (50ng/mL), human Flt-3 ligand (50ng/mL), human thrombopoietin (50ng/mL), and human interleukin-3 (20ng/mL; all cytokines were acquired from PeproTech, Rocky Hill, NJ). Concentrated viral supernatants were used at various MOI to transduce CD34+ HSPCs for 24 hours. These cells were washed, re-plated and cultured under myeloid or erythroid culture conditions; as described by Romero et al (Romero et al., 2013). On day 14 of culture, genomic DNA and/or mRNA was extracted from transduced cells.

### **DDPCR for VCN and % $\beta^{AS3}$ mRNA quantification**

Genomic DNA was extracted using PureLink™ Genomic DNA Mini Kit (Invitrogen, Carlsbad, CA, United States). VCN was calculated by using probes SCD4 (Human Syndecan 4) as a reference and HIV-1 PSI as a target. ddPCR was carried out as described in Urbinati et al (Urbinati et al., 2017). RNeasy Plus Mini Kit (Qiagen, Valencia, CA) was used for RNA extraction followed by reverse transcription as described by Urbinati et al (Urbinati et al., 2017). Probes HBB<sup>TOTAL</sup> as a reference and Hb $\beta^{AS3}$  as a target were used to generate droplets for digital droplet PCR (ddPCR), as described by Hindson et al (Hindson et al., 2011). Droplets were analyzed for absolute quantification of the  $\beta^{AS3}$  gene expression normalized to the total B-globin gene expression.

### ***In vivo* experiment in SCD mouse model**

Bone marrow from eight to twelve-week-old homozygous  $\beta^S/\beta^S$  “Townes” mice (JAX stock #013071) were lineage-depleted using the lineage cell depletion kit from Miltenyi Biotec. Lin<sup>-</sup> cells were pre-stimulated for 24-hours in StemSpan (Stem Cell Technologies, Vancouver, Canada) supplemented with murine stem cell factor (100ng/mL), human interleukin 11 (100ng/ mL), murine interleukin 3 (20 ng/mL), and human FLT-3 ligand (100 ng/mL). Pre-stimulated Lin<sup>-</sup> cells were then transduced at various MOIs to obtain similar VCNs in the bulk cell product or mock transduced. Twenty-four hours later, one to two million transduced cells were delivered by retro-orbital injection after recipient mice (B6-GFP Transgenic; Jackson Laboratory]) were lethally irradiated (1,075 cGy, split in two fractions). A portion of the transduced cells

were cultured for 2-week *in vitro* under myeloid differentiation conditions to determine VCN in the cell product.

Peripheral blood samples were collected at weeks 4 and 16 to measure VCN of engrafted cells by ddPCR, expression of Hb $\beta^{AS3}$  hemoglobin by HPLC, and to determine RBC indices. At week 16, mice were euthanized and BM cells were used to measure engraftment by flow cytometry (GFP+/-), VCN, and expression.

### **High-performance liquid chromatography**

To characterize and quantify hemoglobin tetramers, including human HbS and Hb $\beta^{AS3}$ , and murine HbA and HbF, 1  $\mu$ l of murine peripheral blood was lysed in 25  $\mu$ l hemolysate and incubated at room temperature. Hemolysates were then centrifuged at 500g for 10 minutes at 4°C to remove red blood cell ghost. The lysates were then stored frozen at -80°C and later thawed and processed as described by Urbinati et al (Urbinati et al., 2017).

### **Statistical analysis**

All data are reported as mean  $\pm$  standard deviation of the mean unless otherwise stated. Statistical analyses were performed using GraphPad Prism version 7.0 (GraphPad Software, San Diego, California, USA). The statistical significance between two averages was established using unpaired t-tests. When the statistical significance between three or more averages were evaluated, a one-way ANOVA was applied followed by multiple paired comparisons for normally distributed data (Tukey test). When normality assumption was violated, Mann-Whitney U test was performed for

group-wise comparison instead. Linear regression analyses were used to determine the correlation between VCN and  $\beta^{\text{AS3}}$ -globin RNA transcripts quantities. All statistical tests were two-tailed and a p-value of  $< 0.05$  was deemed significant.

## Tables, Figures and Legends

**Tables 1: Comparison of normalized expression between culture conditions.**

<b>Vector</b>	<b>Erythroid VCN</b>	<b>Erythroid %BAS3/VCN</b>	<b>Myeloid VCN</b>	<b>Myeloid %BAS3/VCN</b>
98.75	3.4 ( $\pm 0.5$ )	6.8 ( $\pm 1.3$ )	3.3 ( $\pm 1.3$ )	Not Detectable
97.5	3.1 ( $\pm 0.8$ )	7.9 ( $\pm 2.7$ )	2.6 ( $\pm 0.7$ )	Not Detectable
95	2.1 ( $\pm 0.4$ )	10.5 ( $\pm 3.3$ )	1.9 ( $\pm 0.7$ )	Not Detectable
90	1.1 ( $\pm 0.1$ )	16.1 ( $\pm 0.7$ )	1.5 (0.7)	Not Detectable
Lenti-BAS3	1.1 ( $\pm 0.2$ )	19.9 ( $\pm 2.4$ )	1.0 ( $\pm 0.3$ )	Not Detectable

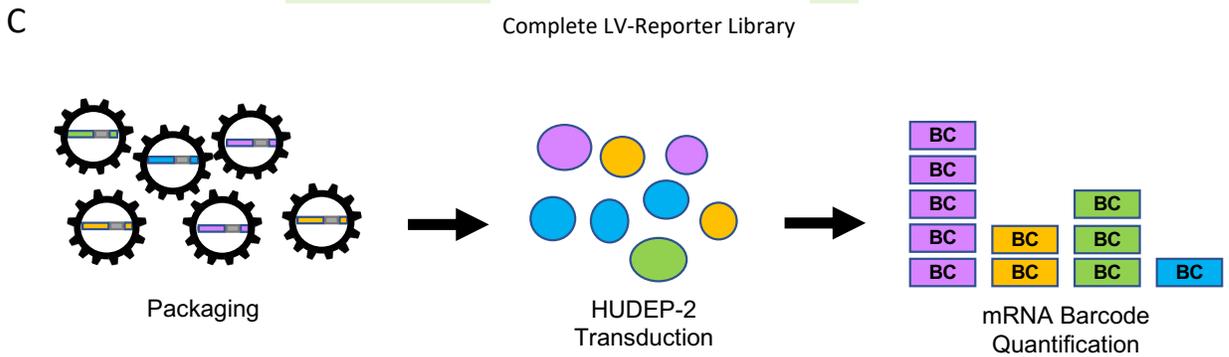
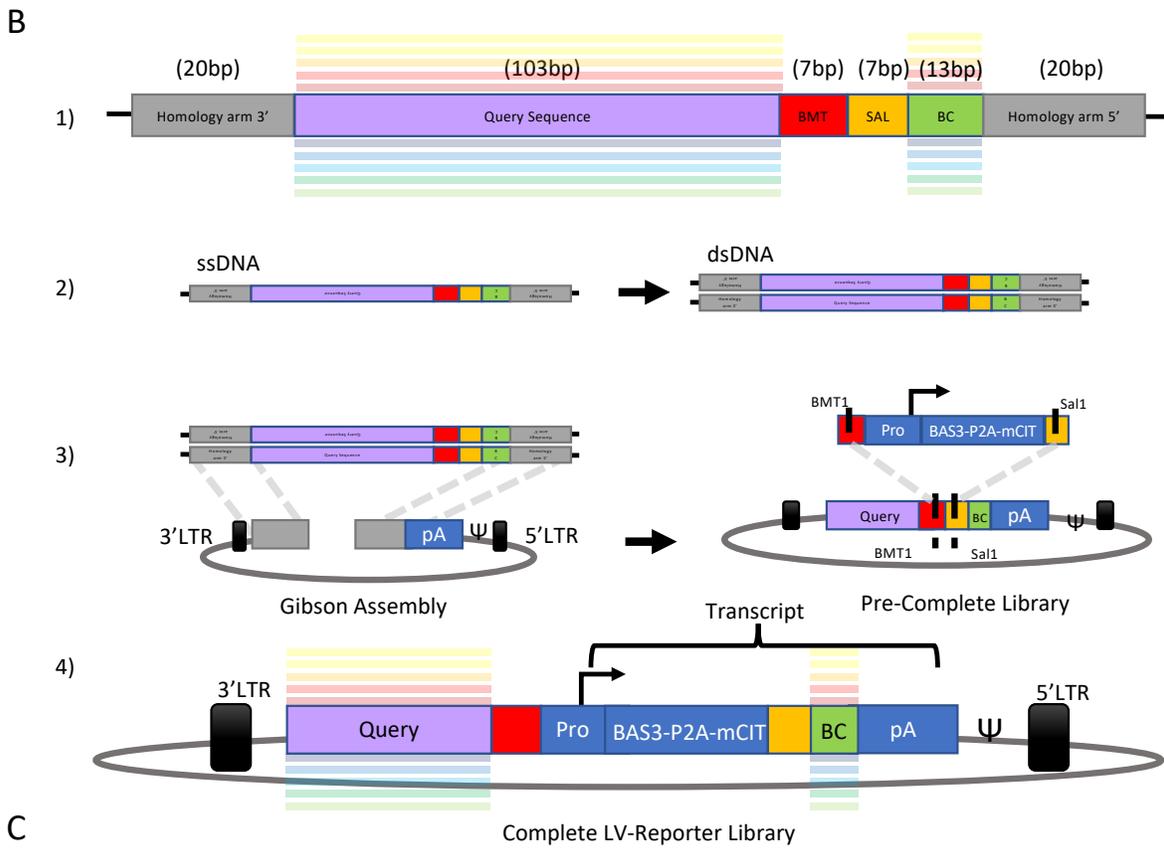
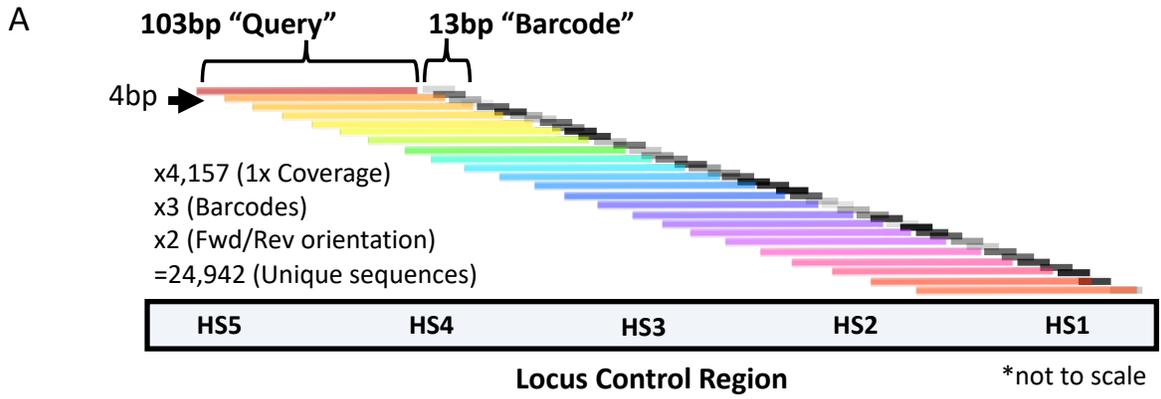
**Tables 2: Sequences of primers used in this study.**

<b>Sequence name</b>	<b>5'→3'</b>
Custom Array Fwd	GTTGGACTTAGGGAACAAAG
Custom Array Rev	ATGTTTTTCTAGGTCTCGAG
1 <sup>st</sup> Strand Primer	AGT AAA ATA TTC AGA AAT AAT TTA AAT
1 <sup>st</sup> Strand Primer (nested)	CAT CAT TGC AAT GAA AAT AAA TGT TTT TTA
2nd Strand Primer	CAA CGA GAA GCG CGA TCA CA

**Figure 1: Overview of LV-MPRA library design and experimental workflow.**

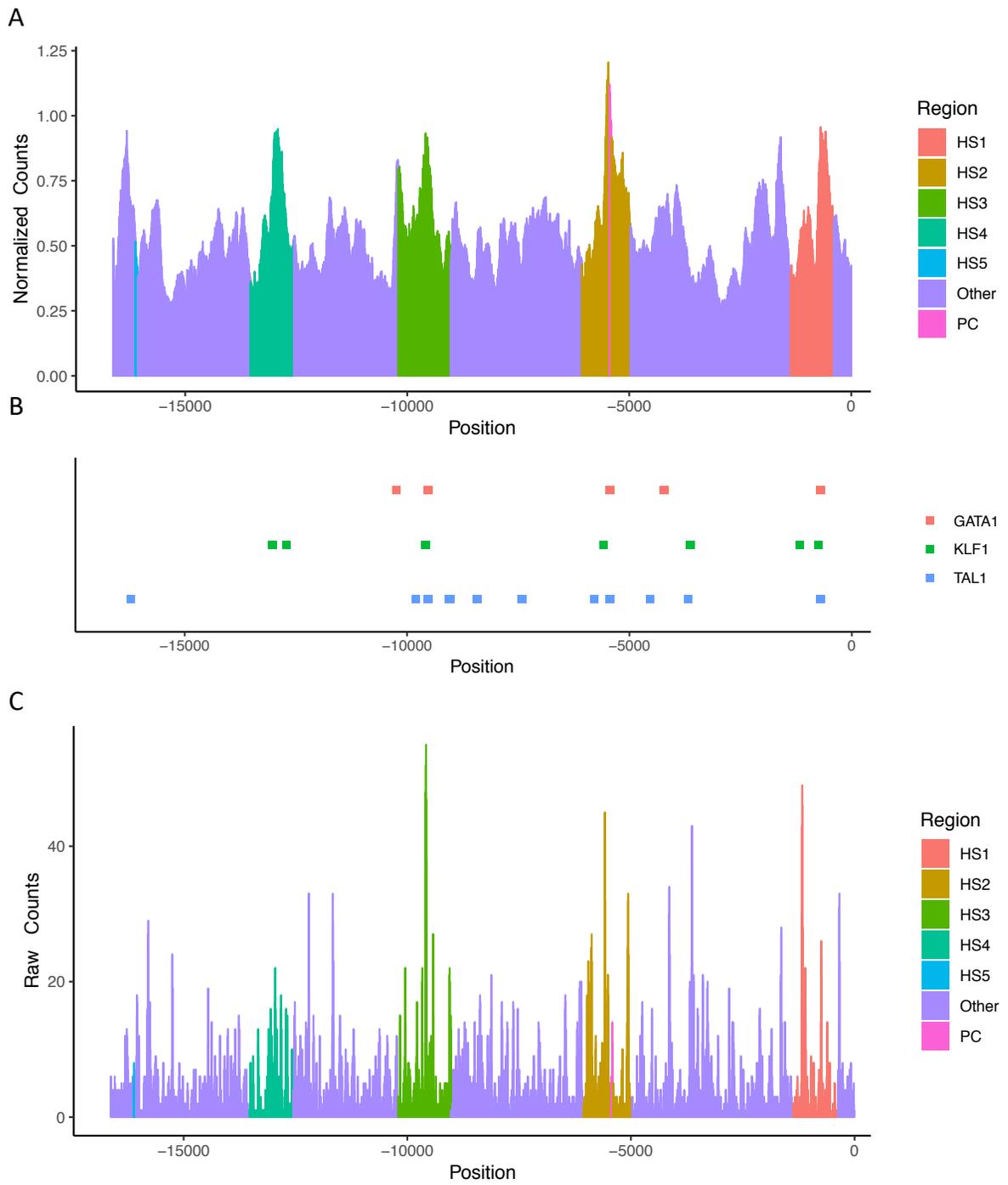
- A) Overlapping 103 bp  $\beta$ -globin Locus Control Region (LCR) sequences were generated with 4 bp tiling. The start of subsequent sequences beginning 4 bp after the start of a preceding sequence. Three unique barcodes (BC) were assigned to each sequence and the entire sequence collection duplicated in reverse orientation and assigned new BCs.
- B) A schematic of a single 170mer is provided.
- I. The 170mer is flanked by 20 bp arms at each end, each possessing homology to the plasmid backbone and are required for downstream cloning. The 103 bp LCR (Query) sequence and 13 bp barcode are separated by *Bmt1* and *Sal1* restriction sites to facilitate downstream cloning.
  - II. A pool of 170mers generated by DNA microarray were converted into dsDNA by primer extension and then PCR amplified.
  - III. The pool of 170 bp dsDNA fragments were joined to lentiviral vector plasmid backbones by Gibson Assembly to ensure 1:1 fusions of 170 bp dsDNA fragments to plasmid. The library was then digested with restriction enzymes and an expression cassette introduced by ligation.
  - IV. The complete library provides placement of the query sequence upstream of the promoter and placement of the BC upstream of a polyadenylation signal to allow for determination of query sequence strength by BC abundance.
- C) The complete LV-Reporter Library was packaged into lentiviral particles and virus used to transduced the erythroid progenitor like cell line, HUDEP-2. The

abundance of cDNA BCs were quantified and normalized to the abundance of plasmid DNA BCs to detect and measure the strength of sequence intrinsic enhancers.



**Figure 2: Sequence intrinsic enhancer map spanning the  $\beta$ -globin Locus Control Region.**

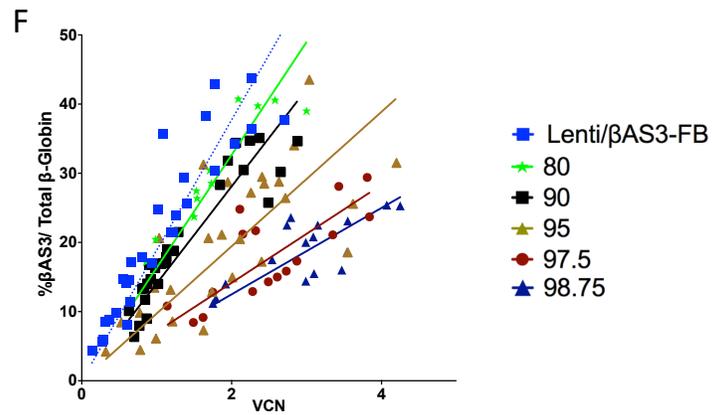
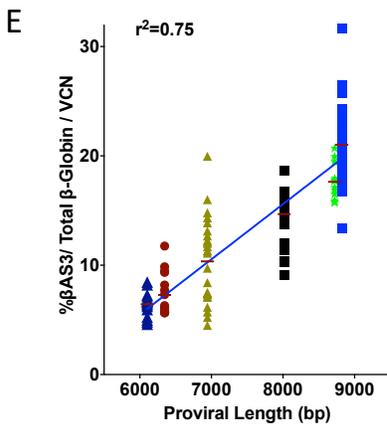
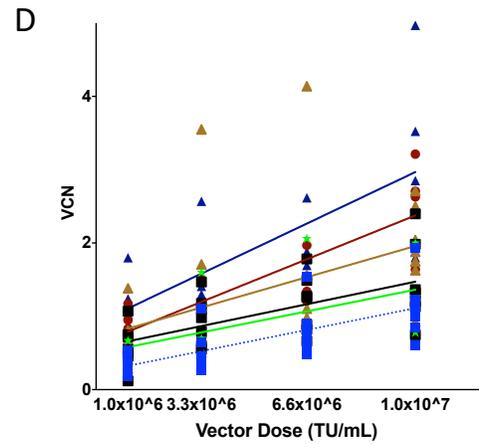
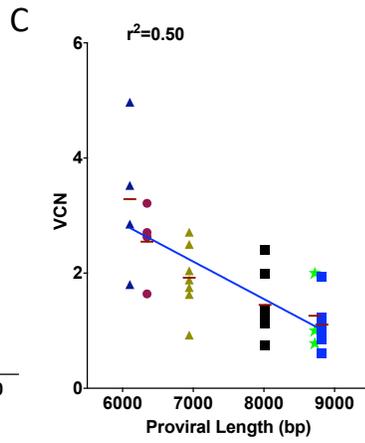
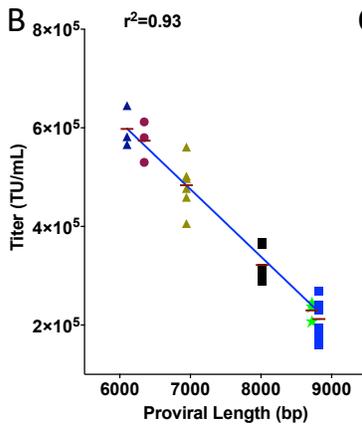
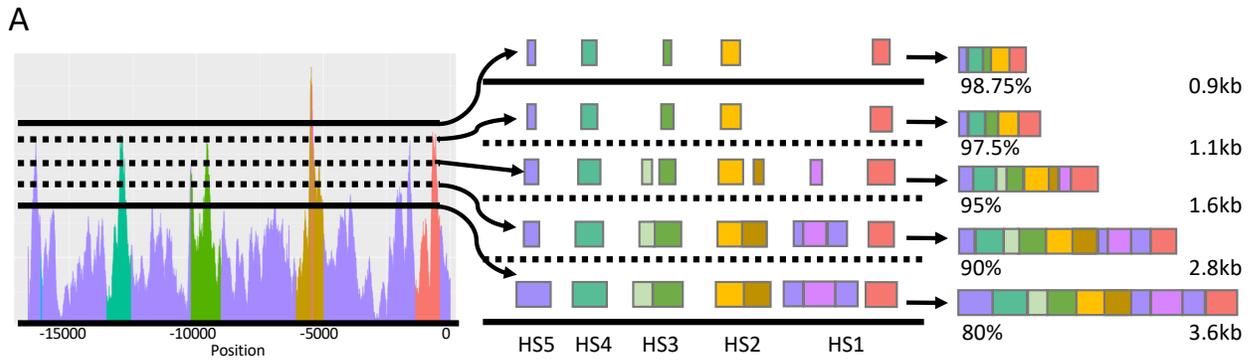
- A) A map of enhancer activity across the LCR is provided. Highlighted regions represent boundaries of LCR hypersensitive sites as defined in previous literature. Track sets available through ENCODE denoting regions of open chromatin and other markings are provided as reference.
- B) A map of important erythroid specific transcription factors.
- C) A map of putative transcription factor binding sites across the LCR. Highlighted regions represent boundaries of LCR hypersensitive sites as defined in previous literature.



### Figure 3: LV-MPRA Guided Therapeutic Vector design and characterization.

- A) Percentile cutoff thresholds were established to identify sequences within a given percentile of expression. Those sequences were then concatenated to produce composite enhancer elements.
- B) Composite enhancers were cloned into the plasmid backbone of a therapeutic lentiviral vector (Lenti/ $\beta$ AS3-FB with LCR DNase hypersensitive sites (HS) 2, 3, and 4 removed), packaged and titered head-to-head, and the quantity of infectious particles plotted as a function of proviral length (bp). Each point in the plot represents an individual 10-cm plate of virus titered on HT-29 cells. Proviral length is defined as sequence length from the beginning of the 5' LTR U3 through the end of the 3' LTR U5. (n)=3-9 per arm.
- C) Human CD34+ hematopoietic stem and progenitor cells (HSPCs) were transduced with constructs at multiplicities of infection (MOI) of 10 ( $1 \times 10^7$  TU/mL) and cultured under myeloid culture conditions to assess infectivity. Vector copy number (VCN) was determined by digital droplet polymerase chain reaction (ddPCR) 14 days after transduction. Each point in the plot represents an individual transduction and the VCN of each transduction is plotted by function of proviral length. (n)=4 per arm.
- D) Human CD34+ HSPCs were transduced at MOIs of 1, 3.3, 6.6 or 10 and cultured under myeloid culture conditions to assess vector infectivity. The VCN of each transduction is plotted by function of vector dose. Slopes represent linear regressions. (n)=12 to 28 per arm.

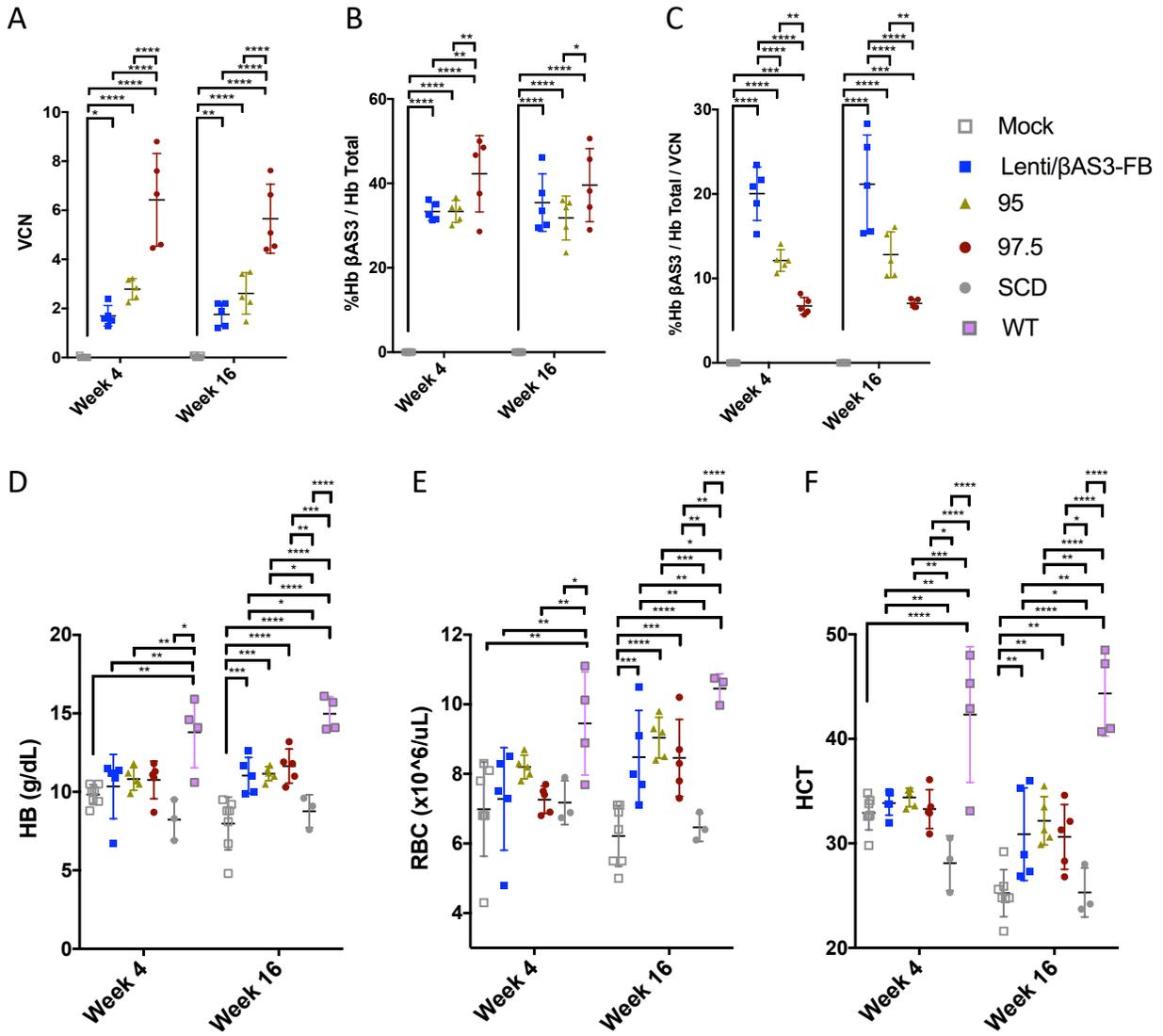
- E) Human CD34+ HSPCs were transduced at MOIs of 1, 3.3, 6.6 or 10 and differentiated under erythroid culture conditions. Percentages of  $\beta$ AS3-globin RNA to total  $\beta$ -globin like RNA were determined by reverse transcription (RT) ddPCR and normalized to VCN. Normalized expression values are plotted as a function of proviral length. (n)=12 to 28 per arm.
- F) Human CD34+ HSPCs were transduced at MOIs of 1, 3.3, 6.6 or 10 and differentiated under erythroid culture conditions. Percentages of  $\beta$ AS3-globin RNA to Total  $\beta$ -globin RNA are plotted by function of their corresponding VCNs. (n)=12 to 28 per arm.



**Figure 4. *In vivo* analysis of peripheral blood from “Townes” mouse model of SCD.**

Peripheral blood (PB) was obtained at weeks 4 and 16. Mice with >97% donor engraftment were analyzed. n=7, Mock; n=5, Lenti/ $\beta$ AS3-FB; (n)=5, 95; (n)=5, 97.5; (n)=3, SCD (“Townes” mouse model of SCD); (n)=3, WT (B6-GFP); \*p<0.05; \*\*p<0.01; \*\*\*p<0.001; \*\*\*\*p<0.0001

- A) Peripheral blood VCN by ddPCR.
- B) Percentages Hb  $\beta$ AS3-globin tetramers in PB lysates measured by high-performance liquid chromatography.
- C) Percentages of Hb  $\beta$ AS3-globin tetramers normalized to PB VCN.
- D) Hemoglobin (HB [g/dL]) levels.
- E) Red Blood Cell (RBC) count ( $\times 10^6$ ).
- F) Hematocrit (HCT) level (percentages).



**Supplementary Figure 1: ~100 bp fragment of LCR DNase HS2 element enhances expression from minimal  $\beta$ -globin promoter.**

- A) Plasmid maps of LV-Reporter (expression driven by only a minimal  $\beta$ -globin promoter) and LV-Reporter (+103 bp HS2).
- B) Bar graph displaying MFI of differentiated HUDEP-2 cells transduced with LV-Reporter or LV-Reporter- (+103 bp HS2). (n)=3 per arm.
- C) Representative FACS histograms. VCN assessed by ddPCR.

A

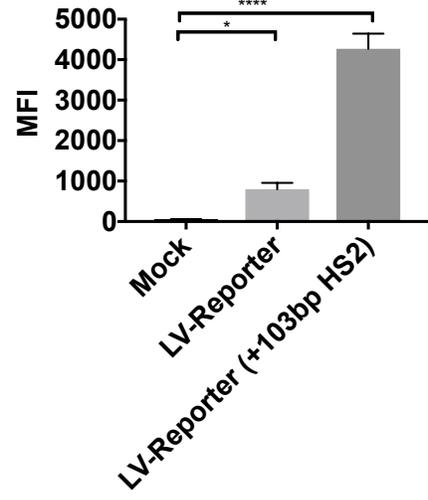
LV-Reporter



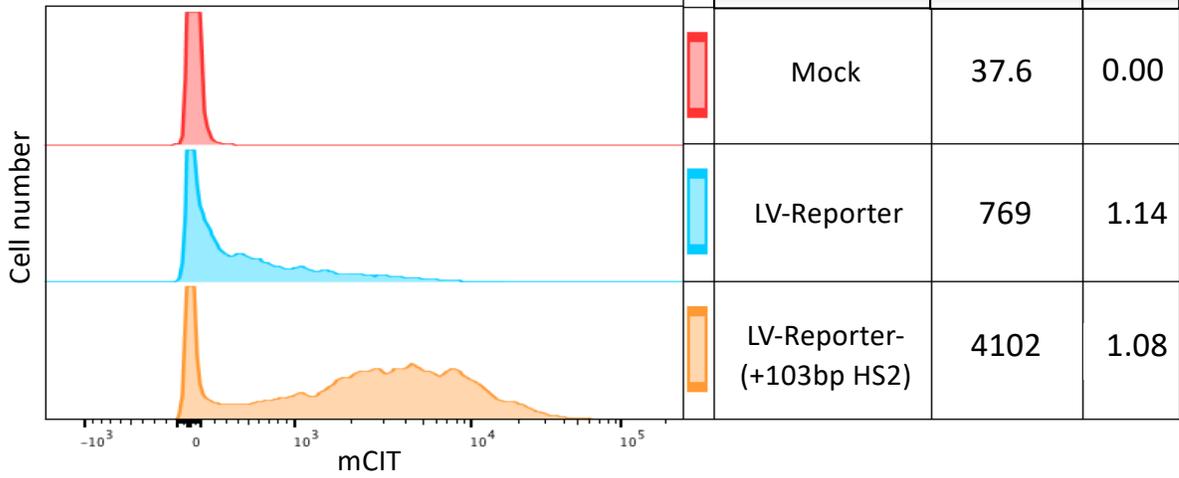
LV-Reporter (+103bp HS2)



B

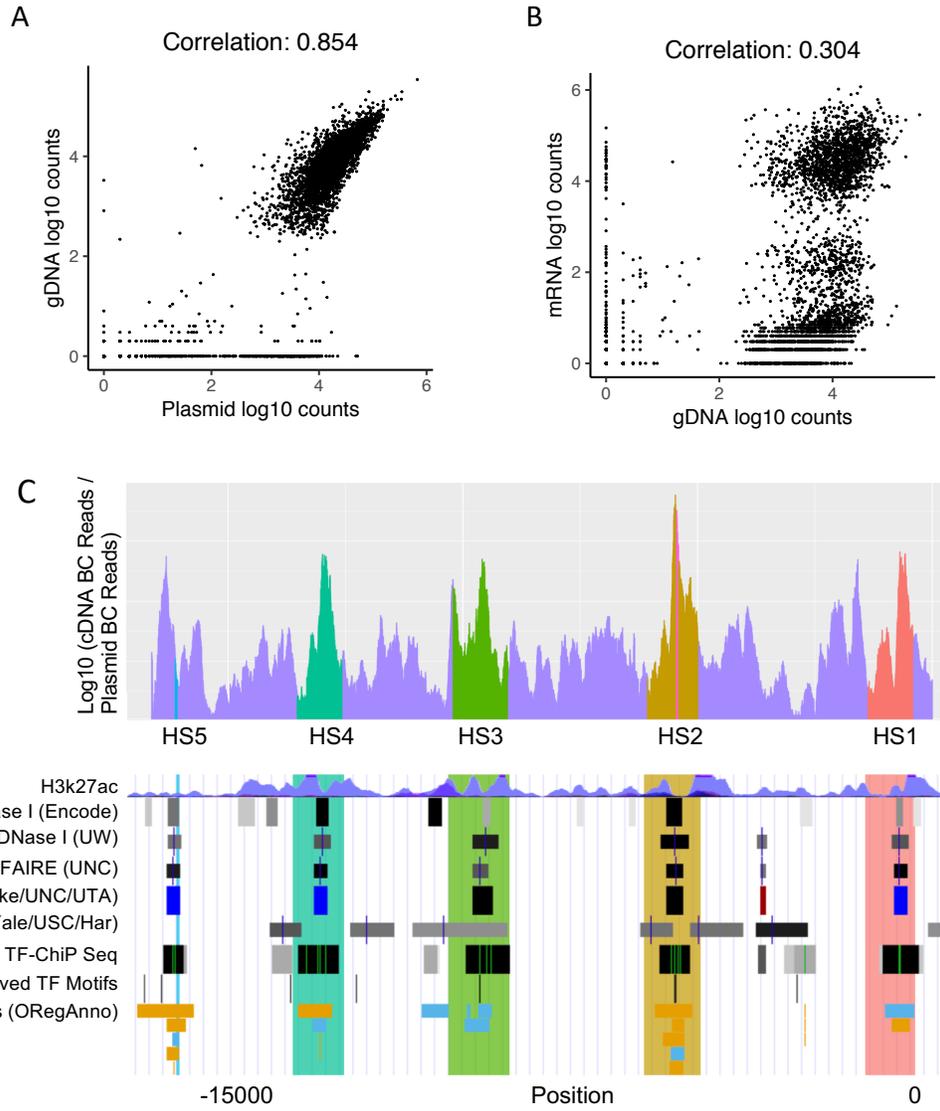


B



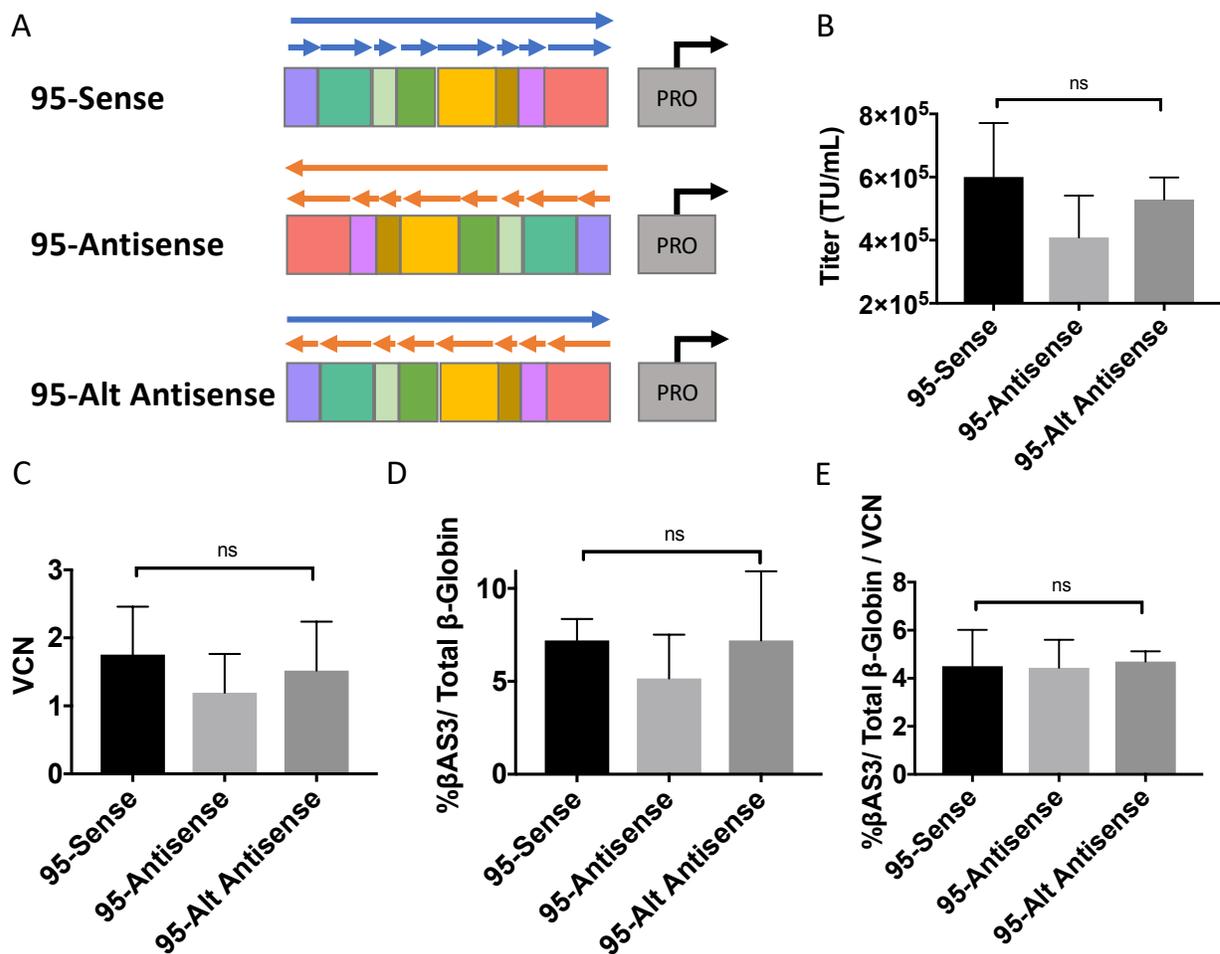
**Supplementary Figure 2: Sequence intrinsic enhancer map spanning the  $\beta$ -globin Locus Control Region aligns with ENCODE.**

- A) Scatter plot showing library transfer efficiency. Each dot represents an individual barcode. Transfer efficiency is plotted as  $\log_{10}$  (Plasmid vs gDNA) BC read count. Correlation between BC counts in plasmid and gDNA is 0.854.
- B) Scatter plot showing expression measurements of “Query” sequences by quantification of mRNA and gDNA BCs derived from HUDEP-2 cells. Expression is plotted as  $\log_{10}$  (RNA vs gDNA) BC read count. Correlation between BC counts in mRNA and gDNA is 0.304.
- C) Map of enhancer activity across the  $\beta$ -globin Locus Control Region is provided. Highlighted regions represent boundaries of LCR hypersensitive sites as defined in previous literature. Track sets available through ENCODE denoting regions of open chromatin and other markings are provided as reference.



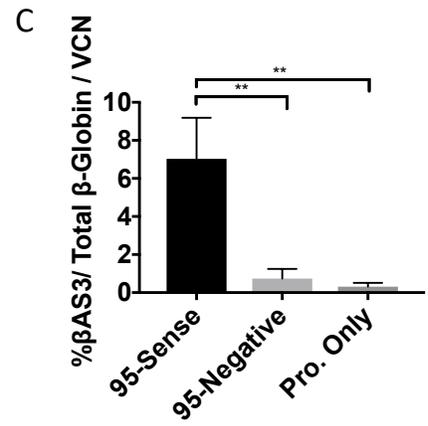
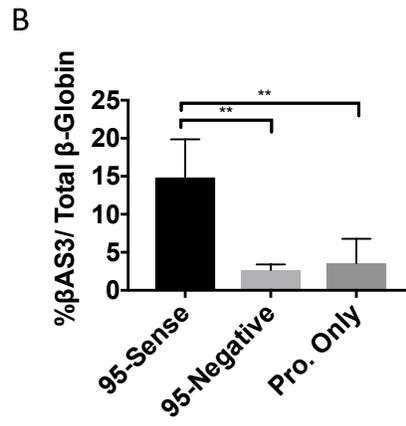
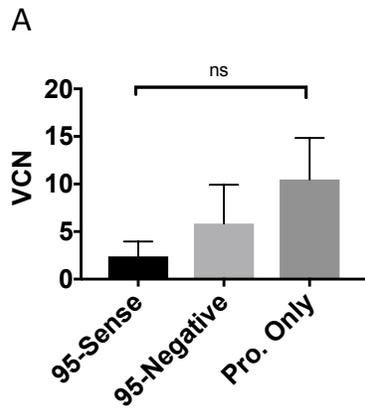
### **Supplementary Figure 3: Determination of optimal enhancer orientation and arrangement.**

- A) Enhancer sequences were generated by imposing a threshold cutoff to identify sequences within the 95<sup>th</sup> percentile of highest expression. Enhancers were then arranged in sense orientation with those associated with HS1 placed closest to the promoter (95-Sense). Enhancers were then arranged in anti-sense orientation with those associated with HS1 placed furthest from the promoter (95-Antisense). Enhancers were again arranged in anti-sense orientation with those associated with HS1 now placed closest to the promoter (95-Alt Antisense). Enhancer configurations were then cloned into Lenti/ $\beta$ AS3-FB (after removal of LCR DNase HS sites 2, 3, and 4).
- B) Quantity of infectious particles packaged and titered in parallel. (n)=3 per arm.
- C) Human CD34<sup>+</sup> HSPCs were transduced with constructs at  $1 \times 10^7$  TU/mL and cultured under myeloid culture conditions for 14-days to assess infectivity. VCN was determined by ddPCR. (n)=3 per arm.
- D) Human CD34<sup>+</sup> HSPCs were transduced in parallel at  $1 \times 10^7$  TU/mL and cultured under erythroid culture conditions for 14-days. Expression of different constructs is presented as percentages of  $\beta$ AS3-globin to total  $\beta$ -globin transcripts.
- E) Expression normalized to VCN.



**Supplementary Figure 4: Sequences within the 95<sup>th</sup> percentile of lowest expression do not enhance expression.**

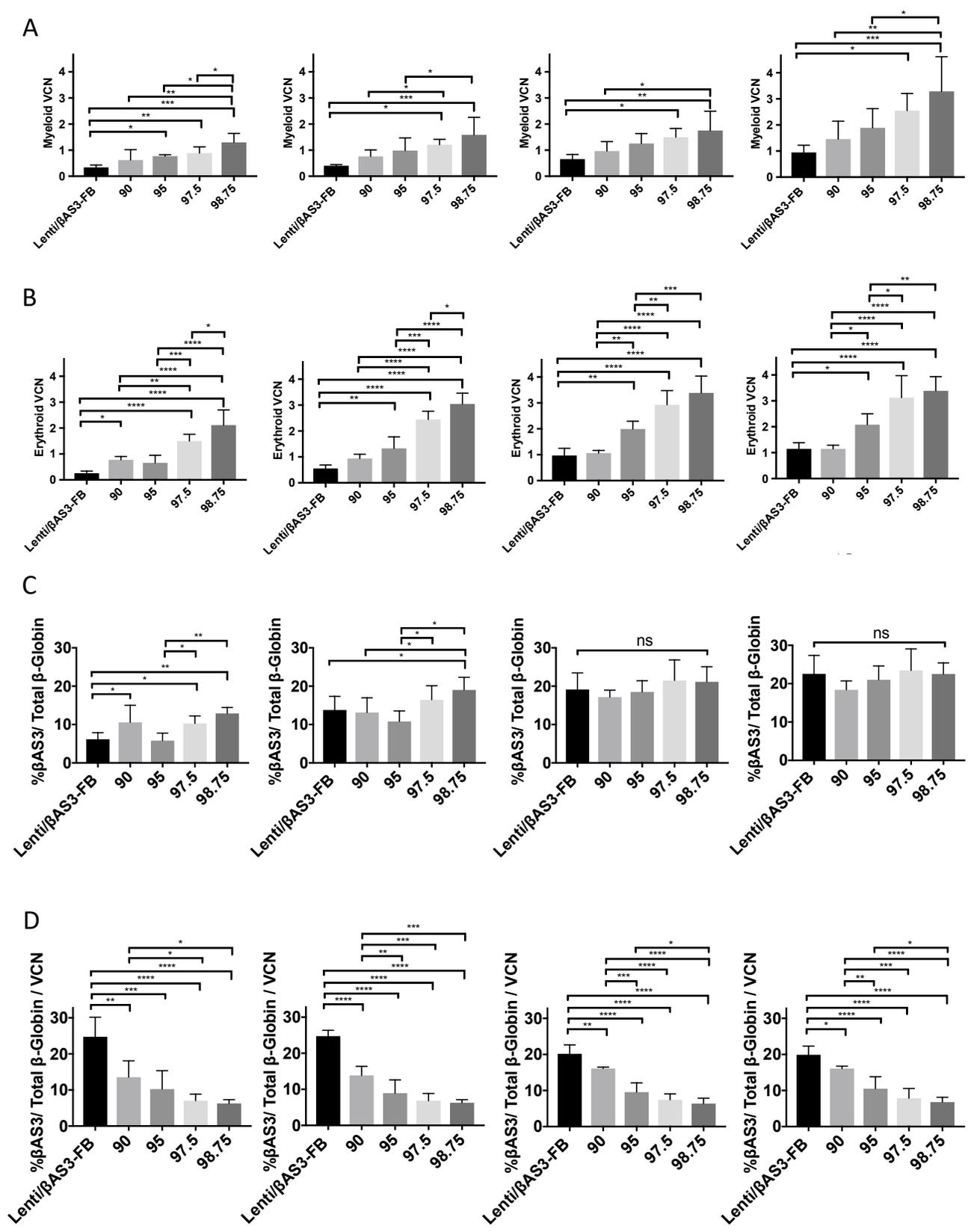
- A) Enhancer sequences for 95-Negative were generated by imposing a threshold cutoff to identify sequences within the 95<sup>th</sup> percentile of lowest expression and cloned into Lenti/ $\beta$ AS3-FB (after removal of LCR DNase HS sites 2, 3, and 4). Human CD34<sup>+</sup> HSPCs were transduced with constructs at  $1 \times 10^7$  TU/mL and cultured under myeloid culture conditions for 14-days to assess infectivity. VCN was determined by ddPCR.
- B) Human CD34<sup>+</sup> HSPCs were transduced in parallel at  $1 \times 10^7$  TU/mL and cultured under erythroid culture conditions for 14-days. Expression of different constructs is presented as percentages of  $\beta$ AS3-globin to total  $\beta$ -globin transcripts.
- C) Expression normalized to VCN.



**Supplementary Figure 5: *In vitro* characterization of vectors designed using LV-MPRA enhancer map.**

Arrangement of bar graphs correspond with transduction conditions. From left to right: 1e6, 3.3e6, 6.6e6 or 1e7 TU/mL, respectively. (n)=4 per arm.

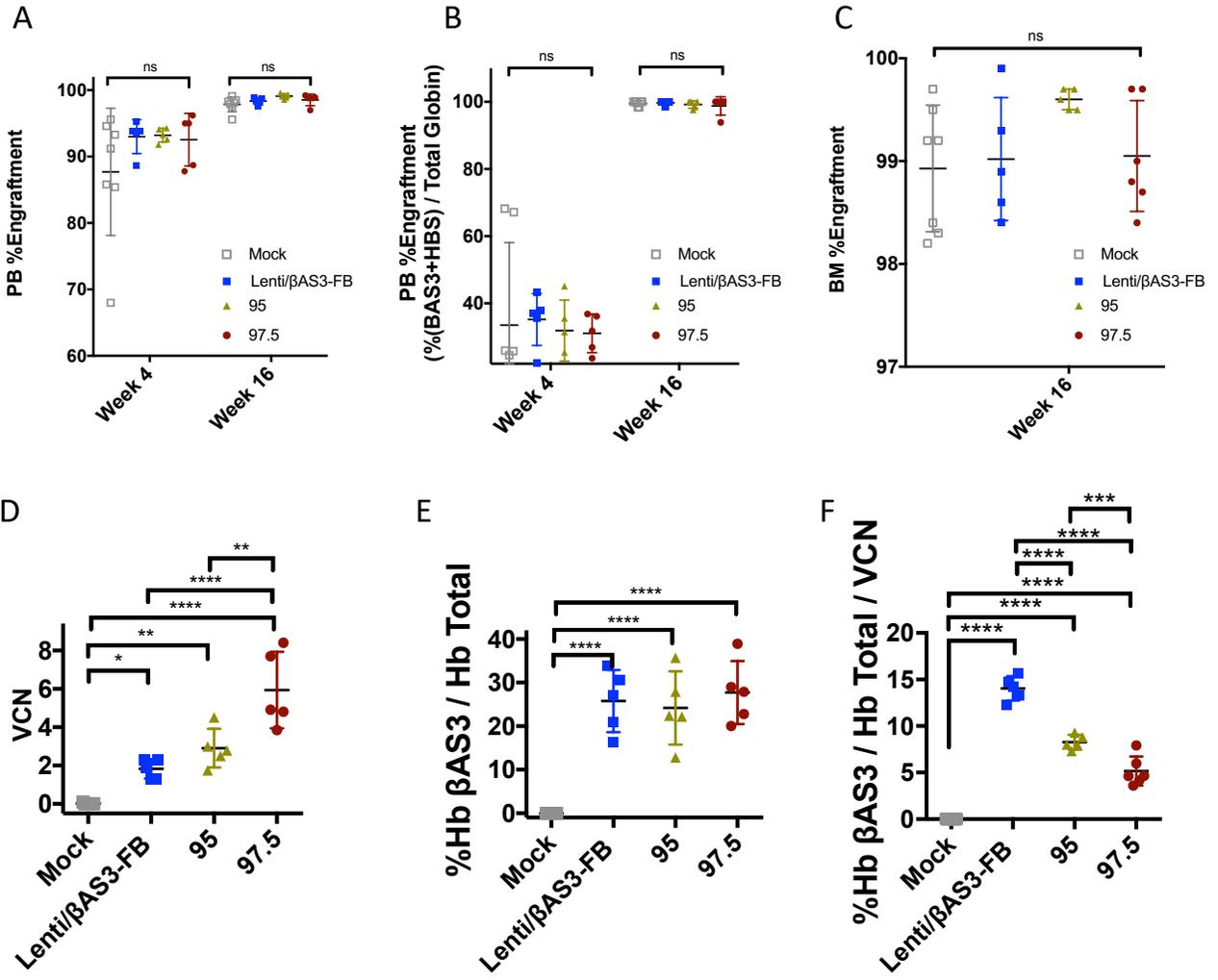
- A) Human CD34+ HSPCs were transduced at various MOIs and cultured under myeloid culture conditions for 14-days to assess infectivity.
- B) Human CD34+ HSPCs were transduced at various MOIs and cultured under erythroid culture conditions for 14-days.
- C) Expression of different constructs is presented as percentages of  $\beta$ AS3-globin to total  $\beta$ -globin transcripts.
- D) Expression normalized to VCN.



**Supplementary Figure 6. Analyses of peripheral blood and bone marrow from “Townes” mouse model of SCD.**

Peripheral blood (PB) was obtained at weeks 4 and 16 and whole bone marrow (BM) was obtained at time of euthanasia (week 16). Mice with >97% donor engraftment were analyzed. n=7, Mock; n=5, Lenti/ $\beta$ AS3-FB; (n)=5, 95; (n)=5, 97.5; \*p<0.05; \*\*p<0.01; \*\*\*p<0.001; \*\*\*\*p<0.0001

- A) PB engraftment levels at weeks 4 and 16 quantified by FACS.
- B) PB engraftment levels at weeks 4 and 16 quantified by HPLC.
- C) BM engraftment levels at week 16 quantified by FACS.
- D) BM VCN by ddPCR.
- E) Percentages  $\beta^{AS3}$ -globin RNA expression of BM determined by RT ddPCR.
- F) Percentages of BM  $\beta^{AS3}$ -globin RNA expression normalized to BM VCN.



## References

1. Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M.P., Baricordi, C., Dionisio, F., Calabria, A., Giannelli, S., Castiello, M.C., et al. (2013). Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome. *Science* 341, 1233151–1233151.
2. Biffi, A., Montini, E., Lorioli, L., Cesani, M., Fumagalli, F., Plati, T., Baldoli, C., Martino, S., Calabria, A., Canale, S., et al. (2013). Lentiviral Hematopoietic Stem Cell Gene Therapy Benefits Metachromatic Leukodystrophy. *Science* 341, 1233158–1233158.
3. Boelens, J.J., Aldenhoven, M., Purtill, D., Ruggeri, A., DeFor, T., Wynn, R., Wraith, E., Cavazzana-Calvo, M., Rovelli, A., Fischer, A., et al. (2013). Outcomes of transplantation using various hematopoietic cell sources in children with Hurler syndrome after myeloablative conditioning. *121*, 3981–3987.
4. Collis, P., Antoniou, M., and Grosveld, F. (1990). Definition of the minimal requirements within the human beta-globin gene and the dominant control region for high level expression. *9*, 233–240.
5. Cooper, A.R., Lill, G.R., Shaw, K., Carbonaro-Sarracino, D.A., Davila, A., Sokolic, R., Candotti, F., Pellegrini, M., and Kohn, D.B. (2017). Cytoreductive conditioning intensity predicts clonal diversity in ADA-SCID retroviral gene therapy patients. *Blood* 129, 2624–2635.
6. Cooper, A.R., Patel, S., Senadheera, S., Plath, K., Kohn, D.B., and Hollis, R.P. (2011). Highly efficient large-scale lentiviral vector concentration by tandem tangential flow filtration. *177*, 1–9.

7. Delzor, A., Dufour, N., Petit, F., Guillermier, M., Houitte, D., Auregan, G., Brouillet, E., Hantraye, P., and Deglon, N. (2012). Restricted transgene expression in the brain with cell-type specific neuronal promoters. *Hum Gene Ther Methods* 23, 242–254.
8. Donnelly, M.L., Luke, G., Mehrotra, A., Li, X., Hughes, L.E., Gani, D., and Ryan, M.D. (2001). Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *J Gen Virol* 82, 1013–1025.
9. Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., and Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. *72*, 8463–8471.
10. Fedosyuk, H., and Peterson, K.R. (2007). Deletion of the human beta-globin LCR 5'HS4 or 5'HS1 differentially affects beta-like globin gene expression in beta-YAC transgenic mice. *Blood Cells Mol. Dis.* 39, 44–55.
11. Fraser, P., Pruzina, S., Antoniou, M., and Grosveld, F. (1993). Each hypersensitive site of the human beta-globin locus control region confers a different developmental pattern of expression on the globin genes. *Genes Dev.* 7, 106–113.
12. Fraser, P., Hurst, J., Collis, P., and Grosveld, F. (1990). DNaseI hypersensitive sites 1, 2 and 3 of the human  $\beta$ -globin dominant control region direct position-independent expression. *18*, 3503–3508.
13. Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018.

14. Grossman, S.R., Zhang, X., Wang, L., Engreitz, J., Melnikov, A., Rogov, P., Tewhey, R., Isakova, A., Deplancke, B., Bernstein, B.E., et al. (2017). Systematic dissection of genomic features determining transcription factor binding and enhancer function. *Pnas* 114, E1291–E1300.
15. Hacein-Bey-Abina, S., Gaspar, H.B., Blondeau, J., Caccavelli, L., Charrier, S., Buckland, K., Picard, C., Six, E., Himoudi, N., Gilmour, K., et al. (2015). Outcomes Following Gene Therapy in Patients With Severe Wiskott-Aldrich Syndrome. *Jama* 313, 1550.
16. Hacein-Bey-Abina, S., Hauer, J., Lim, A., Picard, C., Wang, G.P., Berry, C.C., Martinache, C., Rieux-Laucat, F., Latour, S., Belohradsky, B.H., et al. (2010). Efficacy of Gene Therapy for X-Linked Severe Combined Immunodeficiency. *N. Engl. J. Med.* 363, 355–364.
17. Hindson, B.J., Ness, K.D., Masquelier, D.A., Belgrader, P., Heredia, N.J., Makarewicz, A.J., Bright, I.J., Lucero, M.Y., Hiddessen, A.L., and Legler, T.C. (2011). High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *83*, 8604–8610.
18. Inoue, F., Kircher, M., Martin, B., Cooper, G.M., Witten, D.M., McManus, M.T., Ahituv, N., and Shendure, J. (2017). A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity. *Genome Research* 27, 38–52.
19. Inoue, F., Kreimer, A., Ashuach, T., Ahituv, N., and Yosef, N. (2018). Massively parallel characterization of regulatory dynamics during neural induction. 370452.

20. Kang, Y., Kim, Y.W., Yun, J., Shin, J., and Kim, A. (2015). KLF1 stabilizes GATA-1 and TAL1 occupancy in the human  $\beta$ -globin locus. *Biochimica Et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1849, 282–289.
21. King, A.D., Huang, K., Rubbi, L., Liu, S., Wang, C.-Y., Wang, Y., Pellegrini, M., and Fan, G. (2016). Reversible Regulation of Promoter and Enhancer Histone Landscape by DNA Methylation in Mouse Embryonic Stem Cells. *17*, 289–302.
22. Kreimer, A., Yan, Z., Ahituv, N., and Yosef, N. (2019). Meta-analysis of massive parallel reporter assay enables functional regulatory elements prediction. *bioRxiv* 202002.
23. Kurita, R., Suda, N., Sudo, K., Miharada, K., Hiroyama, T., Miyoshi, H., Tani, K., and Nakamura, Y. (2013). Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS One* 8, e59890.
24. Kwasnieski, J.C., Fiore, C., Chaudhari, H.G., and Cohen, B.A. (2014). High-throughput functional testing of ENCODE segmentation predictions. *24*, 1595–1602.
25. Levasseur, D.N., Ryan, T.M., Pawlik, K.M., and Townes, T.M. (2003). Correction of a mouse model of sickle cell disease: lentiviral/antisickling beta-globin gene transduction of unmobilized, purified hematopoietic stem cells. *Blood* 102, 4312–4319.
26. Levasseur, D.N., Ryan, T.M., Reilly, M.P., McCune, S.L., Asakura, T., and Townes, T.M. (2004). A recombinant human hemoglobin with anti-sickling properties greater than fetal hemoglobin. *J. Biol. Chem.* 279, 27518–27524.

27. Lisowski, L., and Sadelain, M. (2007). Locus control region elements HS1 and HS4 enhance the therapeutic efficacy of globin gene transfer in  $\beta$ -thalassemic mice. *Blood* 110, 4175–4178.
28. Maricque, B.B., Dougherty, J.D., and Cohen, B.A. (2017). A genome-integrated massively parallel reporter assay reveals DNA sequence determinants of cis-regulatory activity in neural cells. 45, e16–e16.
29. May, C., Rivella, S., Callegari, J., Heller, G., Gaensler, K.M., Luzzatto, L., and Sadelain, M. (2000). Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. *Nature* 406, 82–86.
30. Molete, J.M., Petrykowska, H., Bouhassira, E.E., Feng, Y.Q., Miller, W., and Hardison, R.C. (2001). Sequences flanking hypersensitive sites of the beta-globin locus control region are required for synergistic enhancement. *Mol Cell Biol* 21, 2969–2980.
31. Morgan, R.A., Gray, D., Lomova, A., and Kohn, D.B. (2017). Hematopoietic Stem Cell Gene Therapy: Progress and Lessons Learned. *Cell Stem Cell* 21, 574–590.
32. Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A., and Bejerano, G. (2013). Enhancers: five essential questions. *Nat Rev Genet* 14, 288–295.
33. Peterson, K.R., Clegg, C.H., Navas, P.A., Norton, E.J., Kimbrough, T.G., and Stamatoyannopoulos, G. (1996). Effect of deletion of 5'HS3 or 5'HS2 of the human beta-globin locus control region on the developmental regulation of globin gene expression in beta-globin locus yeast artificial chromosome transgenic mice. *Pnas* 93, 6605–6609.

34. Philipsen, S., and Hardison, R.C. (2018). Evolution of hemoglobin loci and their regulatory elements. *Blood Cells Mol. Dis.* 70, 2–12.
35. Pott, S., and Lieb, J.D. (2015). What are super-enhancers? *Nature Genetics* 2014 47:1 47, 8–12.
36. Ribeil, J.-A., Hacein-Bey-Abina, S., Payen, E., Magnani, A., Semeraro, M., Magrin, E., Caccavelli, L., Neven, B., Bourget, P., Nemer, El, W., et al. (2017). Gene Therapy in a Patient with Sickle Cell Disease. *376*, 848–855.
37. Romero, Z., Urbinati, F., Geiger, S., Cooper, A.R., Wherley, J., Kaufman, M.L., Hollis, R.P., de Assin, R.R., Senadheera, S., Sahagian, A., et al. (2013).  $\beta$ -globin gene transfer to human bone marrow for sickle cell disease. *J Clin Invest* 123, 3317–3330.
38. Sadelain, M., Wang, C.H., Antoniou, M., Grosveld, F., and Mulligan, R.C. (1995). Generation of a high-titer retroviral vector capable of expressing high levels of the human beta-globin gene. *92*, 6728–6732.
39. Sessa, M., Lorioli, L., Fumagalli, F., Acquati, S., Redaelli, D., Baldoli, C., Canale, S., Lopez, I.D., Morena, F., Calabria, A., et al. (2016). Lentiviral haemopoietic stem-cell gene therapy in early-onset metachromatic leukodystrophy: an ad-hoc analysis of a non-randomised, open-label, phase 1/2 trial. *The Lancet* 388, 476–487.
40. Smith, R.P., Taher, L., Patwardhan, R.P., Kim, M.J., Inoue, F., Shendure, J., Ovcharenko, I., and Ahituv, N. (2013). Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model. *45*, 1021–1028.

41. Tallack, M.R., Magor, G.W., Dartigues, B., Sun, L., Huang, S., Fittock, J.M., Fry, S.V., Glazov, E.A., Bailey, T.L., and Perkins, A.C. (2012). Novel roles for KLF1 in erythropoiesis revealed by mRNA-seq. *Genome Research* 22, 2385–2398.
42. Tanimoto, K., Sugiura, A., Omori, A., Felsenfeld, G., Engel, J.D., and Fukamizu, A. (2003). Human  $\beta$ -globin locus control region HS5 contains CTCF-and developmental stage-dependent enhancer-blocking activity in erythroid cells. *23*, 8946–8952.
43. The ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *489*, 57–74.
44. Thompson, A.A., Walters, M.C., Kwiatkowski, J., Rasko, J.E.J., Ribeil, J.-A., Hongeng, S., Magrin, E., Schiller, G.J., Payen, E., Semeraro, M., et al. (2018). Gene Therapy in Patients with Transfusion-Dependent beta-Thalassemia. *N. Engl. J. Med.* 378, 1479–1493.
45. Urbinati, F. (2018). Gene Therapy for Sickle Cell Disease :A Lentiviral Vector Comparison Study. *29*, 1153–1166.
46. Urbinati, F., Wherley, J., Geiger, S., Fernandez, B.C., Kaufman, M.L., Cooper, A., Romero, Z., Marchioni, F., Reeves, L., Read, E., et al. (2017). Preclinical studies for a phase 1 clinical trial of autologous hematopoietic stem cell gene therapy for sickle cell disease. *19*, 1096–1112.
47. van Arensbergen, J., FitzPatrick, V.D., de Haas, M., Pagie, L., Sluimer, J., Bussemaker, H.J., and van Steensel, B. (2017). Genome-wide mapping of autonomous promoter activity in human cells. *35*, 145–153.

48. Walters, M.C. (2015). Update of hematopoietic cell transplantation for sickle cell disease. *22*, 227–233.
49. Weber, L., Poletti, V., Magrin, E., Antoniani, C., Martin, S., Bayard, C., Sadek, H., Felix, T., Meneghini, V., Antoniou, M.N., et al. (2018). An Optimized Lentiviral Vector Efficiently Corrects the Human Sickle Cell Disease Phenotype. *Molecular Therapy - Methods & Clinical Development* *10*, 268–280.
50. White, M.A., Myers, C.A., Corbo, J.C., and Cohen, B.A. (2013). Massively parallel in vivo enhancer assay reveals that highly local features determine the cis-regulatory function of ChIP-seq peaks. *Pnas* *110*, 11952–11957.
51. Yanez-Cuna, J.O., Arnold, C.D., Stampfel, G., Boryn, L.M., Gerlach, D., Rath, M., and Stark, A. (2014). Dissection of thousands of cell type-specific enhancers identifies dinucleotide repeat motifs as general enhancer features. *Genome Research* *24*, 1147–1156.
52. Zhan, J., Johnson, I.M., Wielgosz, M., and Nienhuis, A.W. (2016). The identification of hematopoietic-specific regulatory elements for WAS gene expression. *3*, 16077.
53. Zhang, Y., An, L., Yue, F., and Hardison, R.C. (2016). Jointly characterizing epigenetic dynamics across multiple human cell types. *Nucleic Acids Res* *44*, 6721–6731.